ESD-TR-68-128

MTP-75

THE IMPROVEMENT OF DIGITAL HF COMMUNICATION

THROUGH CODING

MAY 1968

ESSI

K. Brayer

Prepared for

AEROSPACE INSTRUMENTATION PROGRAM OFFICE

ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

AD670184

THE IMPROVEMENT OF DIGITAL HF COMMUNICATION
THROUGH CODING

MAY 1968

K. Brayer

Prepared for

AEROSPACE INSTRUMENTATION PROGRAM OFFICE
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

FOREWORD

This report was prepared by the Communications Techniques Department of The MITRE Corporation, Bedford, Massachusetts, under Contract AF 19(628)-5165. The work was directed by the Development Engineering Division under the Aerospace Instrumentation Program Office, Air Force Electronics Systems Division, Laurence G. Hanscom Field, Bedford, Massachusetts. Robert E. Forney served as the Air Force Project Engineer for this program, identifiable as ESD (ESSID) Project 5932, Range Digital Data Transmission Improvement.

REVIEW AND APPROVAL

This technical report has been reviewed and is approved.

OTIS R. HILL, Colonel, USAF
Director of Aerospace Instrumentation
    Program Office

ABSTRACT

In previous papers the technique of error correction of digital data through the use of interleaved cyclic codes and a set of probability functions for the evaluation of error patterns have been presented. In this paper the previous results are extended to a wide range of BCH and symbol codes. A set of simple equations is presented for the description of an interleaved cyclic code and its associated delay, and a method is presented which allows for a significant increase in error rate improvement at a reduction in the delay time introduced into the channel. It is demonstrated that the performance of interleaved cyclic codes is sufficient to correct all types of measured HF error patterns; that, using delay as a basis of comparison, only the total bit interleaving is important in achieving error correction; and that it is possible to get almost 100 percent error correction for delays under 3 seconds for all channel conditions measured.

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

Page

APPENDIX FIGURES: STATISTICS OF TYPICAL TEST RUNS

LIST OF TABLES

PART I

INTERLEAVED CYCLIC CODING

SECTION I

BACKGROUND

In the paper, "Evaluation of Error Correction Block Encoding for High Speed HF Data,"[1] a system of coding was applied to HF digital error patterns collected on the Eastern Test Range.

General characteristics of the observed data error patterns introduced into an HF digital data transmission system were derived and shown to be useful in the evaluation and design of forward error detection and correction using interleaving with cyclic error-correction codes. The technique is applicable to digital data transmission links that require near real-time synchronous operation and low error rates (less than 10^{-5} bit error rate). It is a fundamental requirement that, if a priori knowledge of the message information transmitted is not available at the receive terminal of the transmission link, redundant data bits must be added to the information transmitted to achieve error-detection and correction capability. Thus error-correction capability is obtained at the price of transmission delay time introduced by error-correction implementation and by the addition of redundant data bits to provide the necessary coding structure for each data word transmitted.

The error-pattern data gathered from an operational HF data transmission link was shown to have characteristics that permit the direct classification of the observed error patterns into predominantly random, burst, and periodic error categories. Random-error patterns were characterized by the relatively uniform distribution of errors throughout the observed data. On the other hand, burst-error patterns were characterized by the appearance of clusters of errors in the data stream. Periodic errors were demonstrated to be caused by a high error rate on one of the modem tone channels relative

to the error rate on the other tone channels. Data runs with bit error rates greater than 10^{-3} were predominantly (94 percent) burst-error type runs. Of the data runs with bit error rates less than 10^{-4}, over 72 percent were random-error type runs. The number of random-error data runs that had bit error rates between 10^{-3} and 10^{-4} was approximately equal to that for the burst-error data runs of the same error-rate class. The occurrence of periodic errors increased with increasing bit error rate. For example, 83 percent of the runs in the 10^{-2} error-rate decade contained periodic errors, whereas only 4.5 percent of the runs in the 10^{-5} decade contained periodic errors.

The measure of randomness of observed errors was established by comparison with distributions of known random-error data derived from a computer program that was designed to generate random-error pattern data using a random-number generator. The magnitude of the area between the observed error data distribution and the corresponding known random-error data run was used to evaluate the degree of randomness of observed data errors.

Six typical data runs were selected from a total of 151 observed data runs. These data runs were then analyzed using the characteristic distribution functions of the error patterns. Three of the data runs were selected as random-error data runs, and the other three were selected as burst-error data runs. Error correction without interleaving was evaluated for the random data runs, and it was found that the best performance is achieved using powerful random-error-correcting codes such as a Bose-Chaudhuri (255, 123, 19) code. The three burst-error runs were used to evaluate performance of error correction with and without interleaving. Interleaving used with a modified Golay error-correcting code gave results superior to those obtained using this code without interleaving. Thus a valid method was established to permit the quantitative evaluation of interleaver

- 4 -

performance in terms of effective randomization of burst-error patterns. Criteria were also established for proper design of the interleaver to minimize the effect of periodic errors. There were, however, many questions left unanswered. Some of these questions, which will be answered here, are: How do codes of rates other than one-half perform? How is delay related to performance? Are long symbol codes useful on HF error patterns?

SECTION II

DATA DESCRIPTION

As in previous papers[1,2,3] the approach used in code evaluation will be that of using typical runs. A typical run (Table I) is typical in the sense that it represents the type of error pattern frequently found at some error rate.

Table I

Typical Test Runs

Test Run No.	Data Rate (bits/sec)	Length (min)	Average Bit Error Rate	Source
12	2400	90	2.7×10^{-3}	Antigua to Cape Kennedy
24	2400	90	2.8×10^{-5}	Philco AN/USC-12 modem
309	2400	10	4.3×10^{-3}	Antigua to Ascension Looped
339	2400	10	1.2×10^{-2}	Kineplex TE-216 modem

A description of the TE-216 can be found in Reference [1] and of the AN/USC-12 in Reference [2]. Further information is available from manufacturers' catalogs. The test procedure can be found in Reference [1]. A detailed description of these test runs is in the Appendix. It is demonstrated in the Appendix that test run 12 is characterized by bursts of errors but that

- 7 -

these bursts contain random and periodic errors and are not high-density bursts. Test run 24 consists of short, high-density bursts. Test runs 309 and 339 contain bursts in which the error rate within the burst is high and the intervals between bursts (guard spaces) are short. These test runs also contain periodic errors. At the conclusion of Part II of this paper, the coding system which is found to be best will be evaluated against all HF data previously reported on[3]

SECTION III

PERFORMANCE OF BOSE-CHAUDHURI-HOCQUENHEM CODES

Bose-Chaudhuri-Hocquenhem (BCH) codes[4,5] are a class of cyclic, random-error correcting codes. They are described by the notation (n, k, e) where k information bits are encoded together to obtain n total bits, and of these n bits e bit errors can be corrected. The term "code rate" is defined by the relation

$$\text{code rate} = k/n \qquad (1)$$

The codes are treated here as interleaved codes. In interleaved codes the code words of length n are created by encoding information bits I to generate parity bits P according to the following:

$$I_{(k-1)m+1+s} \quad \cdots \quad I_{2m+1+s} \quad I_{m+1+s} \quad I_{1+s}$$

$$P_{(n-k-1)m+1+s} \cdots P_{2m+1+s} \quad P_{m+1+s} \quad P_{1+s} \qquad (2)$$

$$s = 0, 1, \ldots (m-1)$$

The interleaving parameter is m. If m is one, there is no interleaving. The following example will demonstrate the value of interleaving.

Example: If the modified Golay code is chosen, the defining parameters are (24, 12, 3). If m is 1 , then 12 consecutive information bits are encoded to create the 24-bit code word. In this code word if three errors are introduced they are corrected. If six consecutive errors are introduced there is no correction and actually some error generation. If, however, m were 2, then 12 bits chosen as either the odd or the even bits from the first 24 bits are encoded together to get an interleaved block of two code words. If six

consecutive errors occur in the channel they will fall three each in the two distinct code words, and if there are no other errors in the new total block of two words, the errors will be corrected.

From the point of view of consecutive errors, interleaving redefines the code as (mn, mk, me). The objective of interleaving is to spread a burst of errors into the guard space which immediately follows it. Since the guard space has few errors, a code block is created where the actual number of errors in the code block is less than "me," and, since the bursts have been spread out, error correction can be achieved. The penalty for the improvement due to interleaving is in time delay. In an (n, k, e) code all n bits must be received before all corrections can be performed; thus there is a delay of n bit times at the data rate. The delay in the interleaved system is 2 mk information bit times. The general equation[2] for delay is

$$\text{Delay} = \frac{2 \text{ mk}}{\text{Information Data Rate}}. \tag{3}$$

If the information data rate is 1200 bits/sec and the code rate is 0.5 then the channel rate will be 2400 bits/sec. Thus we find that

$$\text{Information Data Rate} = (\text{code rate}) (\text{channel data rate}); \tag{4}$$

also

$$\text{mk} = \frac{\text{mnk}}{\text{n}} = \text{mn (code rate)}. \tag{5}$$

Thus an alternate equation to (3) which will be useful later is

$$\text{Delay} = \frac{2\ mn}{\text{Channel Data Rate}} = \frac{mn}{1/2(\text{Channel Data Rate})}. \qquad (6)$$

Using Equation (4) it is seen that delay has been normalized to a half-rate code.

The performance of these codes is described as percent of input errors corrected. However, for ease of display, improvement factor is defined as:

$$\text{Improvement Factor} = \frac{1}{1 - \dfrac{\%\ \text{of Errors Corrected}}{100}} \qquad (7)$$

In order to evaluate the performance of BCH codes a computer program was developed to take the actual error patterns measured[1,2,3,6,7] and find the percent of errors corrected for various codes. The results are presented in Figures 1 – 4, showing improvement factor as a function of code rate. All curves are identified by a doublet of which the first number is n and the second is normalized delay in seconds. The channel data rate is fixed at 2400 bits/sec since all error patterns were measured at this data rate. The codes evaluated are BCH codes where

$$n = 2^Q - 1 \qquad 4 \leq Q \leq 8. \qquad (8)$$

A horizontal line with arrows on it indicates that 100 percent of the errors were corrected (infinite improvement). The curves are plotted as smooth curves, but values were obtained only for codes which exist.

- 11 -

Figure 1a. BCH Code Performance,
Test Run 12, n = 15

Figure 1b. BCH Code Performance,
Test Run 12, n = 31

- 12 -

Figure 1c. BCH Code Performance,
Test Run 12, n = 63

Figure 1d. BCH Code Performance,
Test Run 12, n = 127

- 13 -
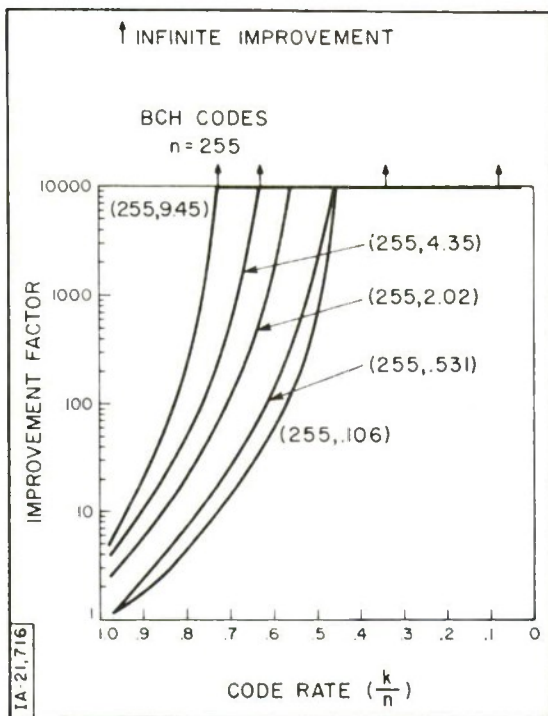
Figure 1e. BCH Code Performance, Test Run 12, n = 255
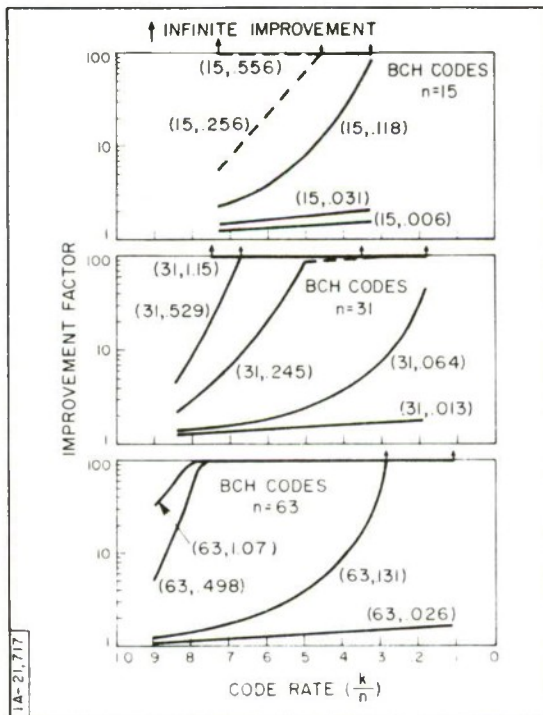
Figure 2a. BCH Code Performance, Test Run 24, n = 15; 31; 63
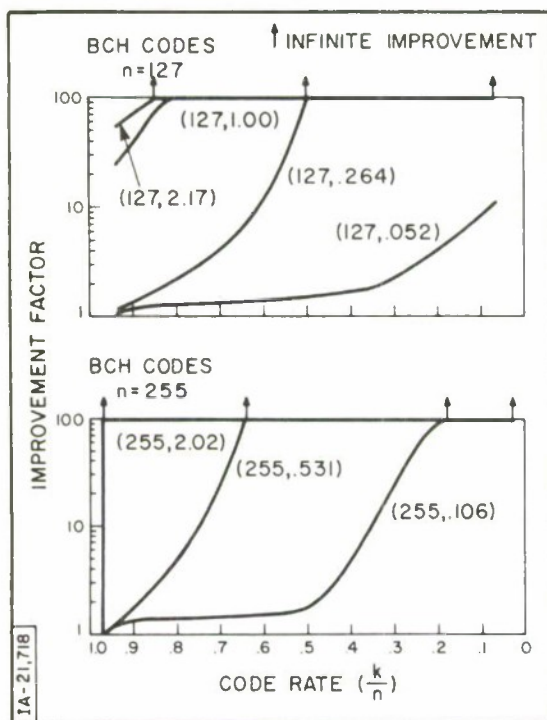
Figure 2b.   BCH Code Performance,
Test Run 24, n = 127; 255
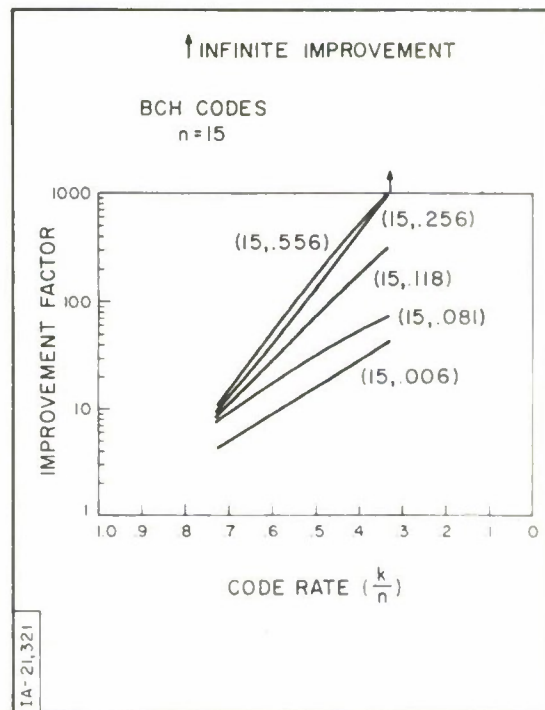
Figure 3a.   BCH Code Performance,
Test Run 309, n = 15
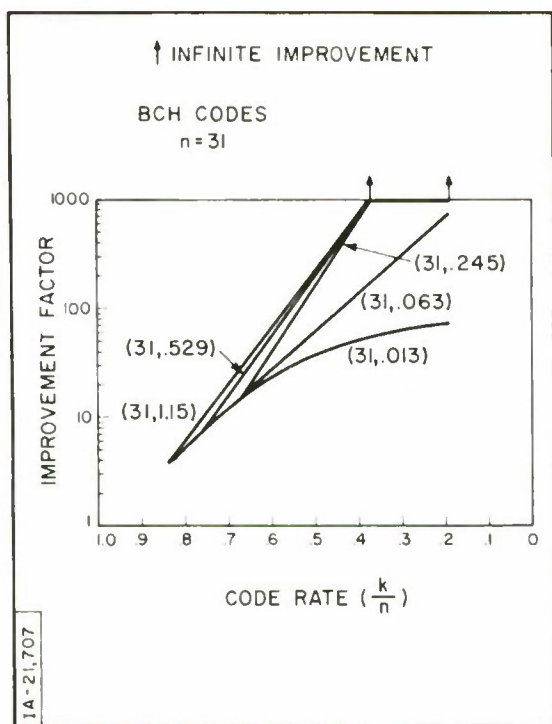
Figure 3b. BCH Code Performance,
Test Run 309, n = 31
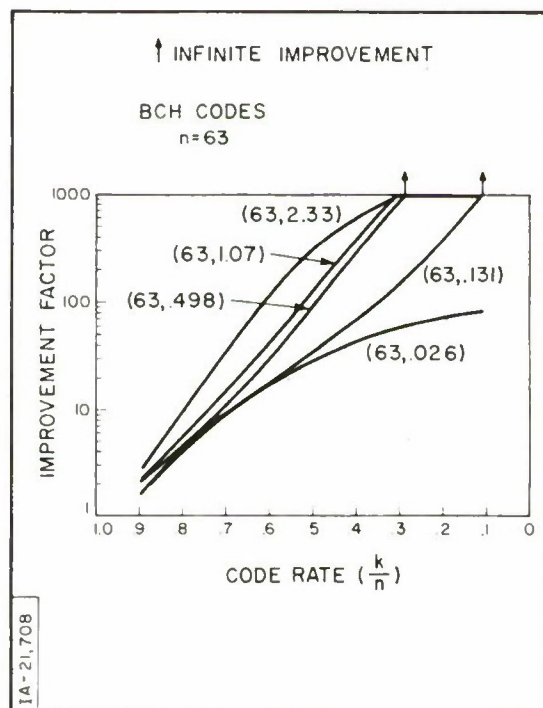
Figure 3c. BCH Code Performance,
Test Run 309, n = 63
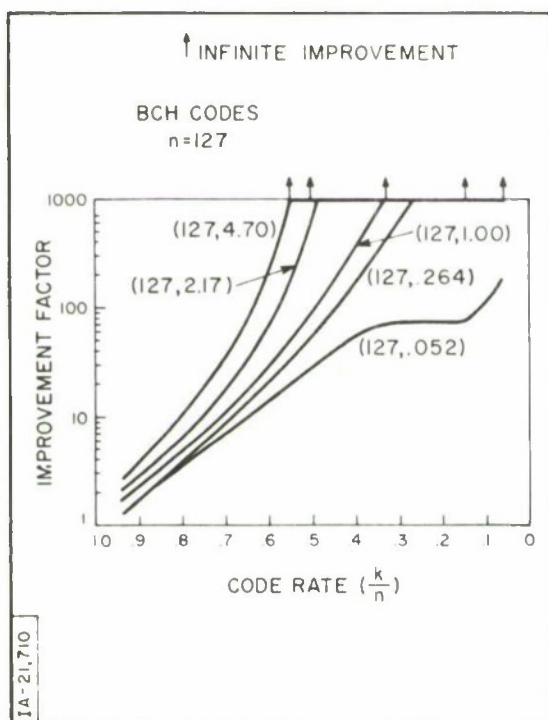
Figure 3d. BCH Code Performance,
Test Run 309, n = 127
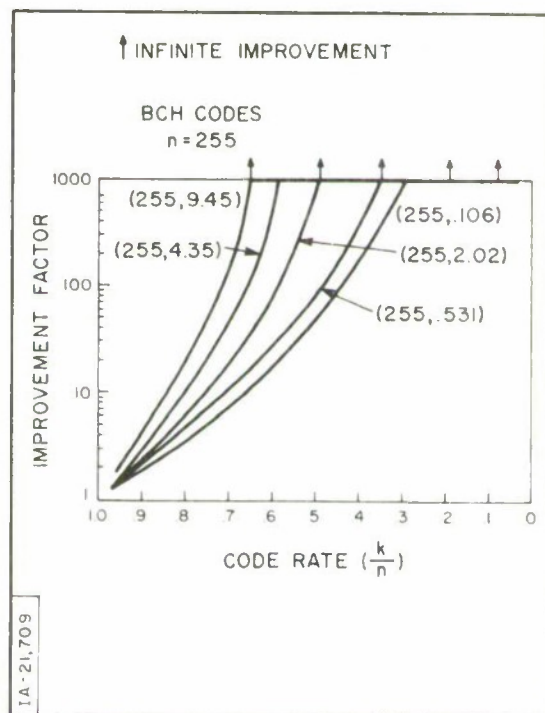
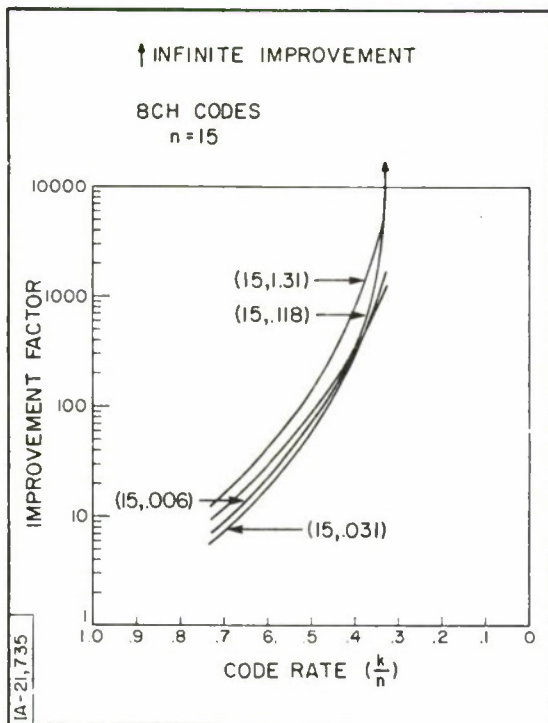Figure 3e. BCH Code Performance,
Test Run 309, n = 255

- 17 -
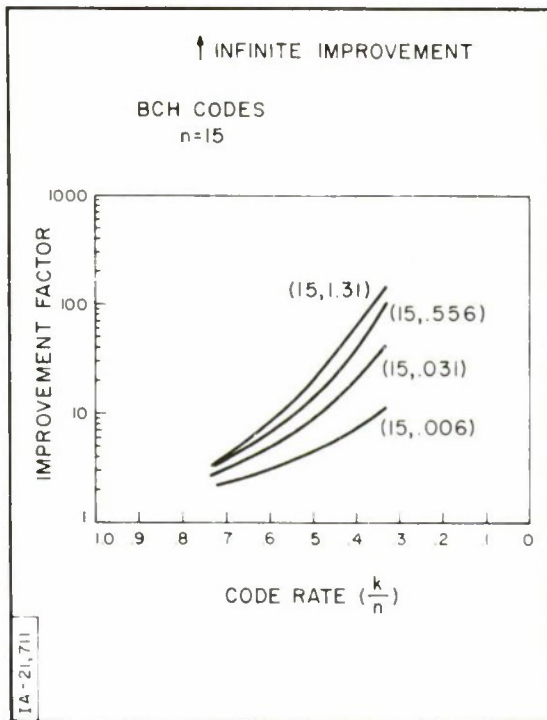
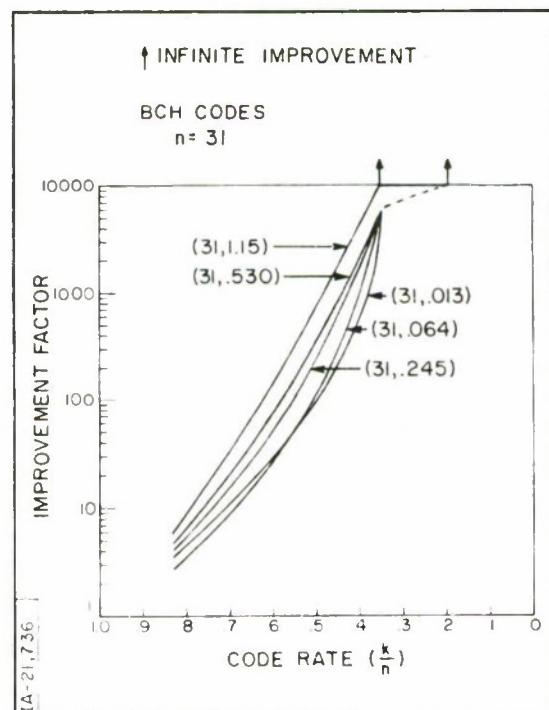Figure 4a.   BCH Code Performance, Test Run 339, n = 15
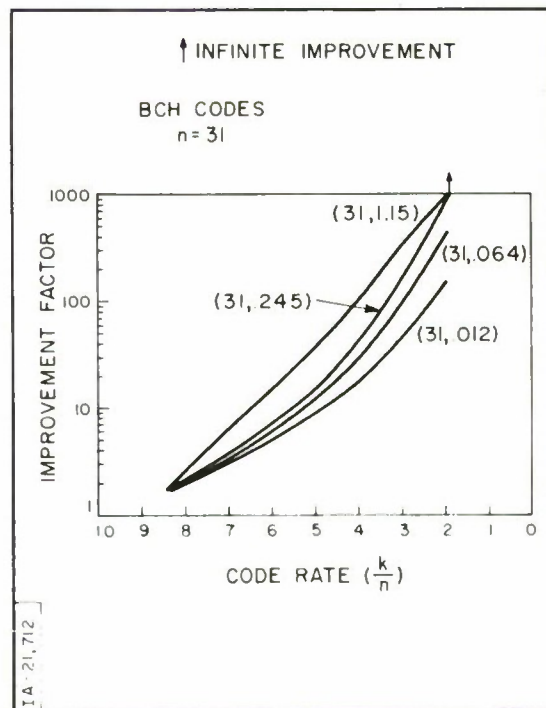
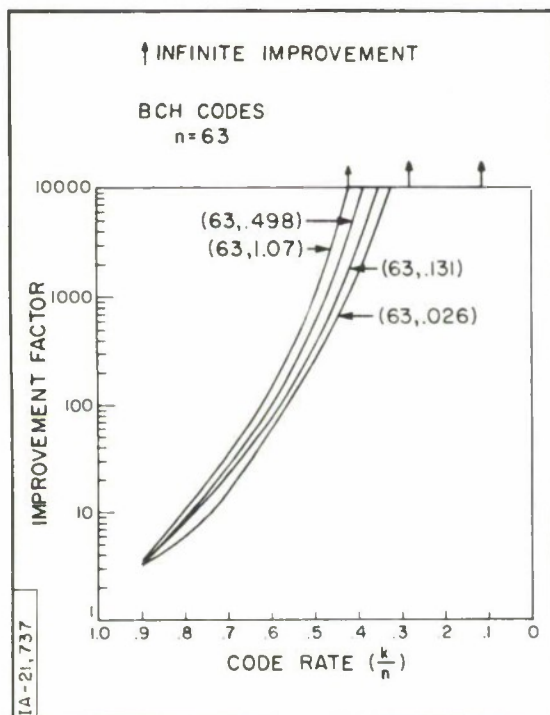Figure 4b.   BCH Code Performance, Test Run 339, n = 31
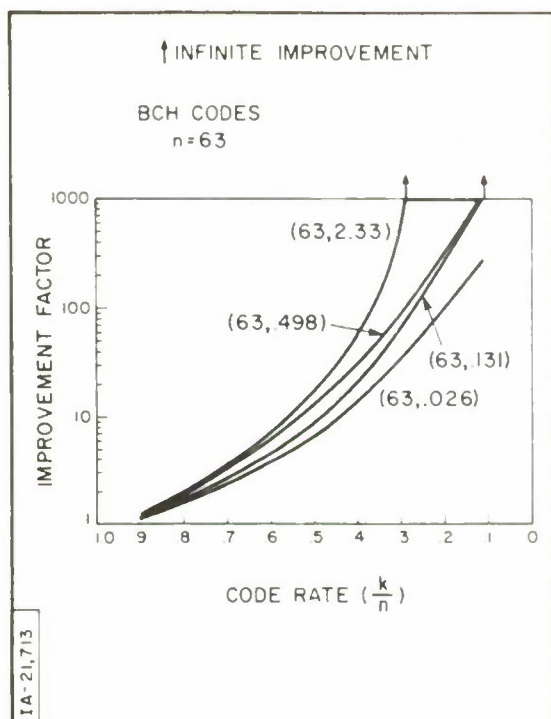
Figure 4c. BCH Code Performance,
Test Run 339, n = 63
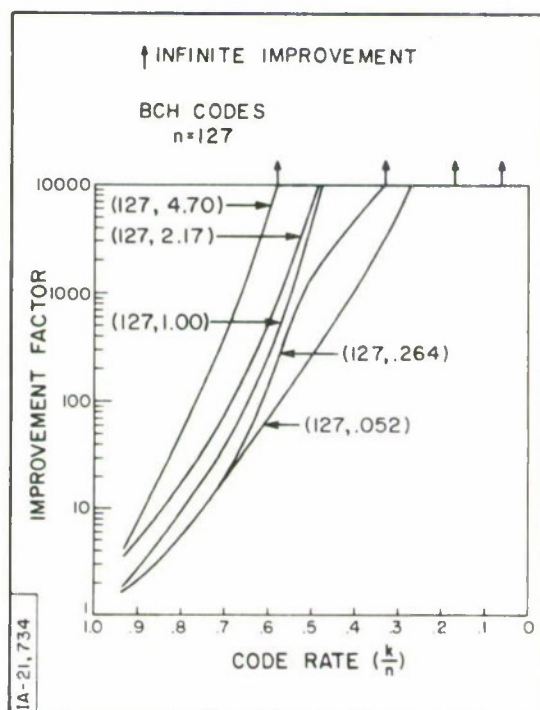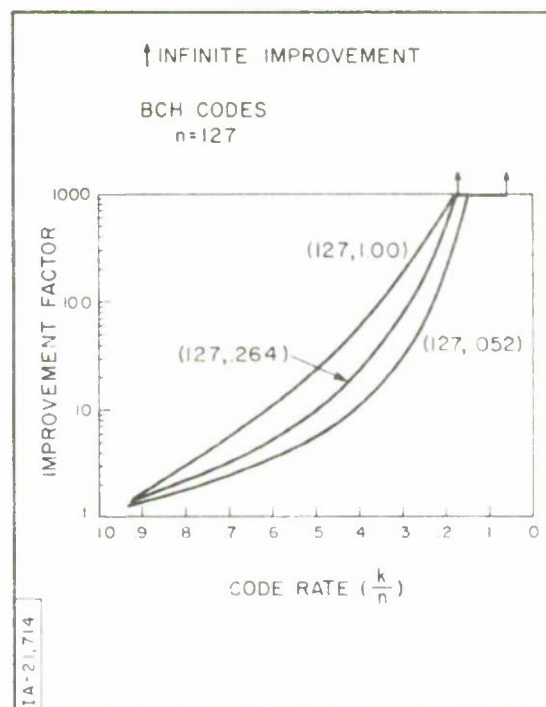

Figure 4d. BCH Code Performance,
Test Run 339, n = 127

- 19 -
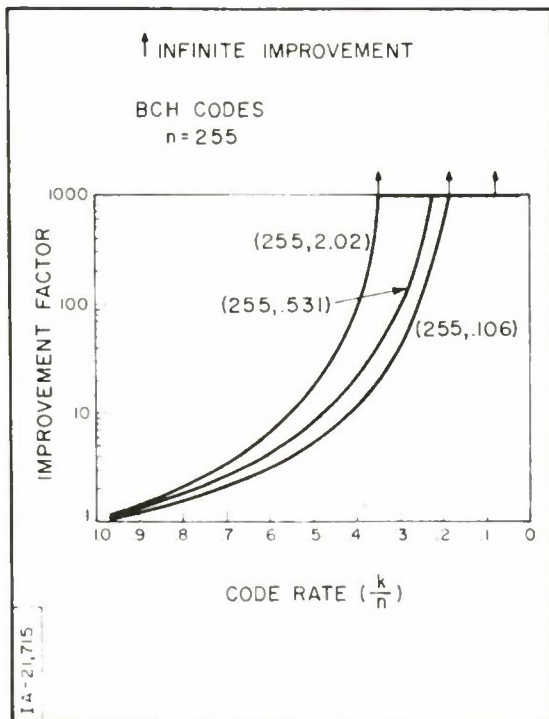
Figure 4e. BCH Code Performance, Test Run 339, n = 255

Considering the set of curves for any one test run, it is interesting to note that for the same delay all codes are approximately equal in performance. On test run 309 the performance for n = 127 at a delay of 2.17 (m = 41) is almost identical to that for n = 255 at a delay of 2.02 (m = 19). Thus it can be concluded that performance is a function of the product mn and not of m or n individually, if delay is considered the common ground upon which two or more codes are compared. The values of m used are 1, 5, 19, 41, and 89, although in some cases the upper values are not presented owing to lack of significant improvement in performance. In all cases it is possible to get 100 percent correction if the code rate is sufficiently reduced (i.e., an increasing portion of the channel is allocated to parity). It can be concluded that, at a reasonable penalty in delay, the error rate in HF digital communication can be substantially reduced by the interleaved BCH code technique.

- 20 -

SECTION IV

PERFORMANCE OF P^M SYMBOL CODES

Symbol codes[8,9] are structured in a fashion similar to BCH codes. In a symbol code a group of bits is taken together as a symbol, a set of information symbols is encoded together to get a block, and some portion of the symbols can be corrected. The parameter M is used to designate the number of bits/symbol. There is a total of $N = 2^M - 1$ symbols/block of which e symbols can be corrected if there are 2e parity symbols. The code rate is as given below:

$$\text{code rate} = \frac{\text{information symbols}}{\text{total symbols}} = \frac{N - 2e}{N}. \tag{9}$$

All previous relations are the same when expressed in symbols, except delay:

$$\text{Delay} = \frac{2 \cdot 2 \text{ mMN}}{\text{Channel Data Rate}} = \frac{2 \text{ mMN}}{1/2(\text{Channel Data Rate})} \tag{10}$$

The factor of two is introduced into the delay because of the extremely complicated calculations[9] necessary in decoding and is already accounted for on the curves (Figures 5 - 8). The performance of this system for interleaving of m = 1 and m = 10 is presented in Figures 5 - 8.

The performance of these codes is, even for the high delays, inferior to that of the BCH codes. The reason lies in the structure of the codes. If m = 1, M = 8, N = 255 is considered for run 339, the total bits/block is 2040. At a code rate near 0.5, a total of 64 symbol errors or 512 consecutive bit

- 21 -
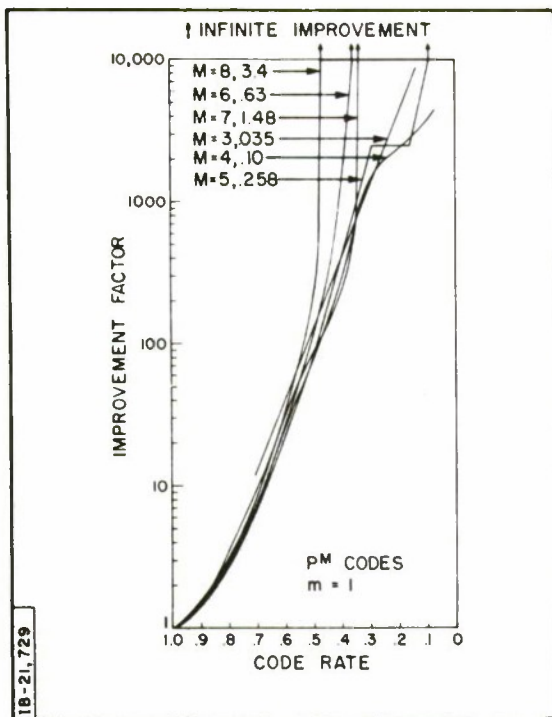
Figure 5a. P^M Code Performance, Test Run 12, m = 1
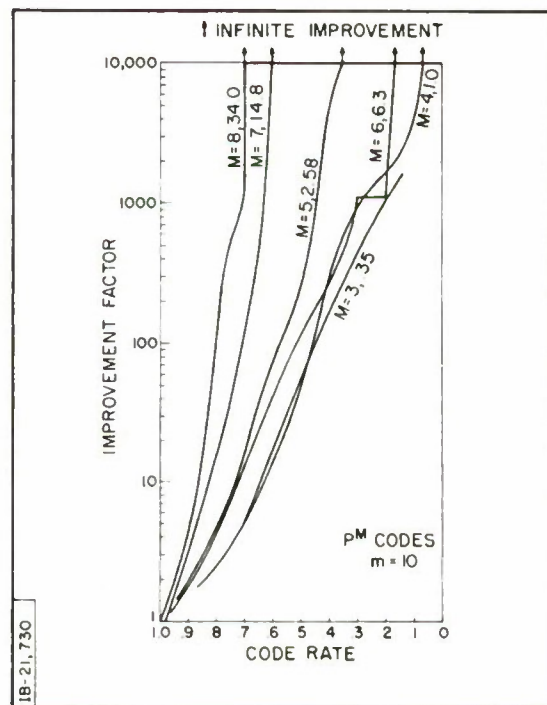
Figure 5b. P^M Code Performance, Test Run 12, m = 10

- 22 -
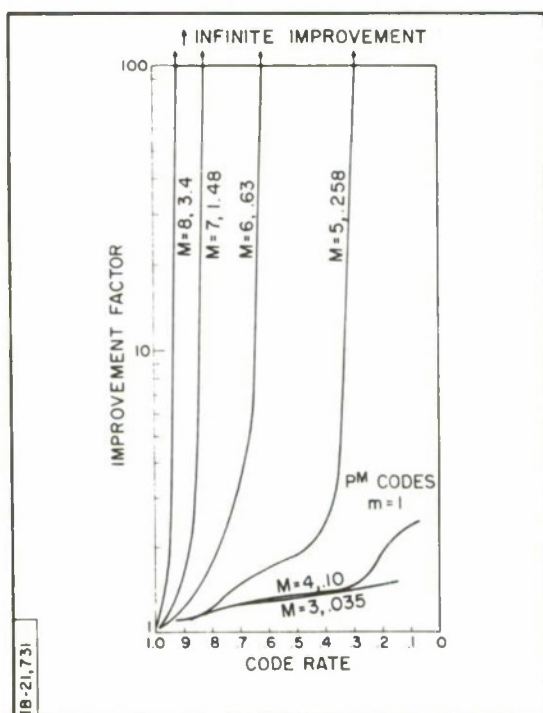
Figure 6a. P^M Code Performance, Test Run 24, m = 1
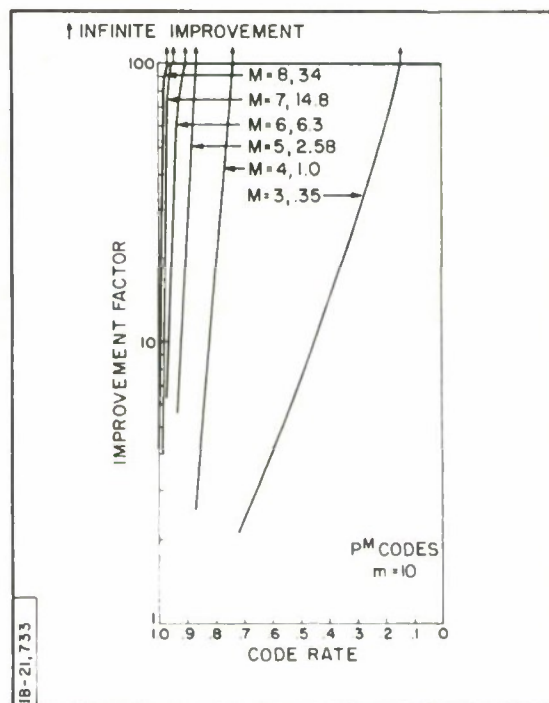
Figure 6b. P^M Code Performance, Test Run 24, m = 10

- 23 -
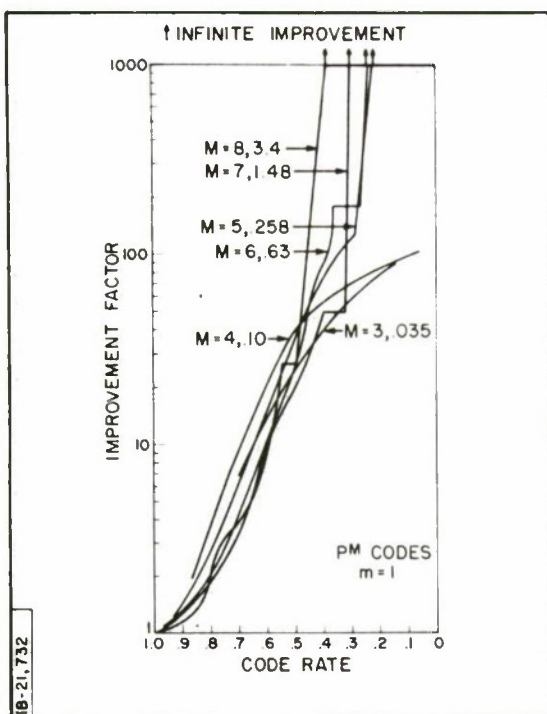
Figure 7a.   P^M Code Performance,
Test Run 309, m = 1
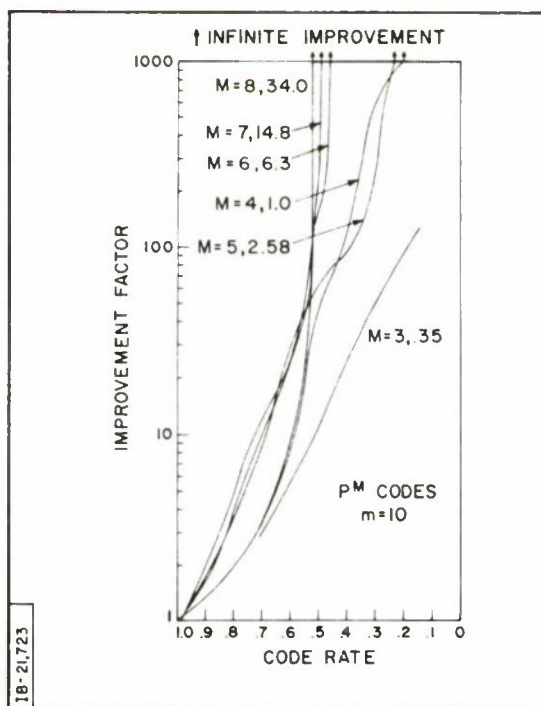
Figure 7b.   P^M Code Performance,
Test Run 309, m = 10

- 24 -
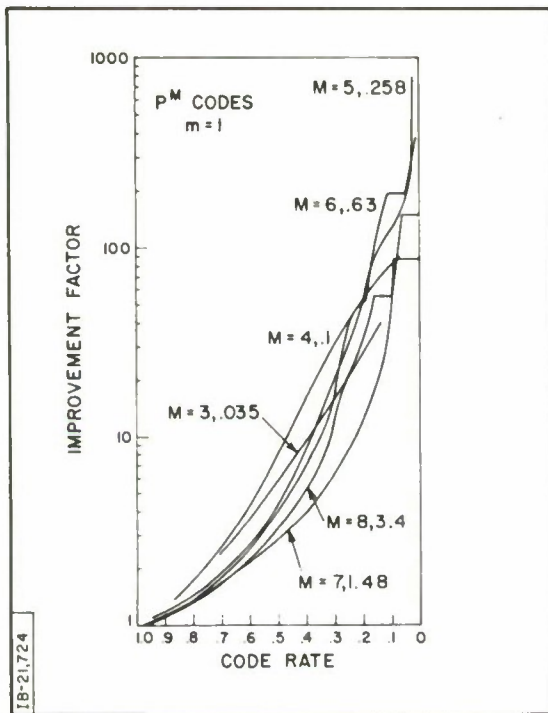
Figure 8a. P^M Code Performance,
Test Run 339, m = 1
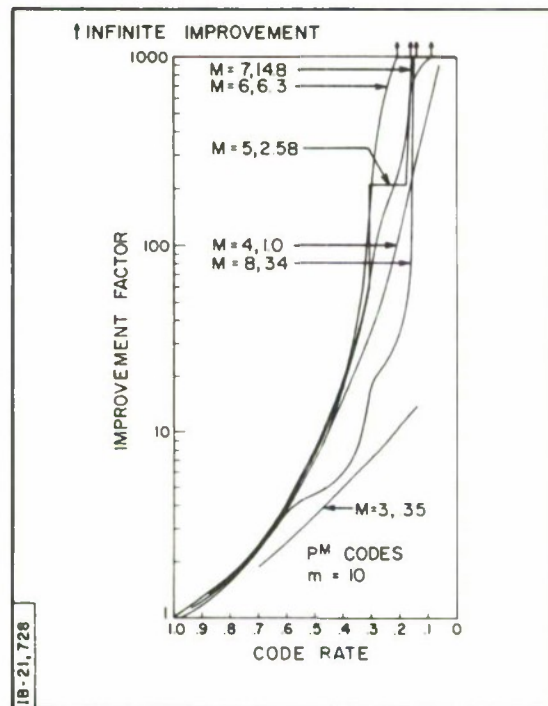
Figure 8b. P^M Code Performance,
Test Run 339, m = 10

- 25 -

errors can be corrected. The bursts are not solid, however, but are spread out, and frequently more than 64 symbols are in error even though less than 512 bits/block are in error. Thus, although the delay is 3.4 seconds the improvement factor is less than 5. Since the bursts in run 339 are spread out, a code (BCH) that corrects bit errors is superior. On run 24, which has the solid bursts that the symbol code prefers to correct, the bursts are short and the symbol codes (with and without interleaving) do as well as or better than the BCH codes. Unfortunately, very little of the observed data is similar to run 24. In another study[9] these codes were shown to work excellently with long bursts of the type observed in test run 24.

SECTION V

CONCLUSIONS

The objective of Part I has been to examine the performance of easily implemented cyclic codes with interleaving on typical samples of error patterns that occur in HF communication. It has been found that when codes of equivalent error-correcting capability are used, the overall performance is governed only by the product of code length (n) and interleaver length (m) (i.e., the delay introduced into the system). Further, it has been demonstrated that the performance of symbol codes (i.e., codes with a greater dense-burst correcting capability) is inferior to that of binary codes for a given amount of delay. The reason for this is that the bursts which occur in the channel are generally low-density, diffuse bursts.

The results obtained in this work are extended to all previously described HF data[3] and to a more sophisticated coding system[10] in Part II of this paper.

PART II

TANDEM INTERLEAVED CYCLIC CODING

SECTION I

INTRODUCTION

The National Range Division (NRD) presently operates high-speed (2400 bits/sec) HF communication links to transmit data back to the Cape Kennedy complex from down-range stations such as Ascension Island. Because of the high error rates normally experienced on these HF channels, error correction must be implemented to improve communication. Previous papers[1,2,3] have described the performance of simple cyclic codes on HF error patterns measured upon these circuits. It has been demonstrated that using half-rate codes (50 percent of the bits transmitted are parity) and delays from five to ten seconds it is possible to get correction of 90 percent of the errors that occur. The method of encoding used has been interleaved cyclic coding. In interleaved cyclic encoding, information bits I are encoded together to get parity bits P according to the following relations:

$$
\begin{array}{cccc}
I_{(k-1)m+1+s} & \cdots \; I_{2m+1+s} & I_{m+1+s} & I_{1+s} \\
\\
P_{(n-k-1)m+1+s} & \cdots \; P_{2m+1+s} & P_{m+1+s} & P_{1+s}
\end{array}
\tag{11}
$$

where $s = 0, 1, \ldots (m-1)$

      n is the number of bits in a code word

      k is the number of information bits

      m is the interleaving parameter

As has been demonstrated in Part I, the effect of interleaving is to redefine the code as an (mn, mk, me) code where the number of correctable

random errors e is multiplied by m, thus increasing the randomization of the errors by placing errors that occur in the same burst into different code words.

In this paper the interleaving technique will be extended to the tandem (concatenated) coding approach presented by Forney[10] (and block diagrammed in Figure 9). In the tandem coding approach, information from the data source is given to an outer coder which performs interleaved encoding and passes the encoded information and its associated parity to the inner coder, which performs the same functions on the total data stream. The inner coder provides the data to the communication equipment that constitutes the transmit portion of the channel. Errors are introduced by the action of the HF channel and demodulation failures, and are corrected by the decoding procedures. This system will be simulated and demonstrated to be superior to using a single interleaved code.
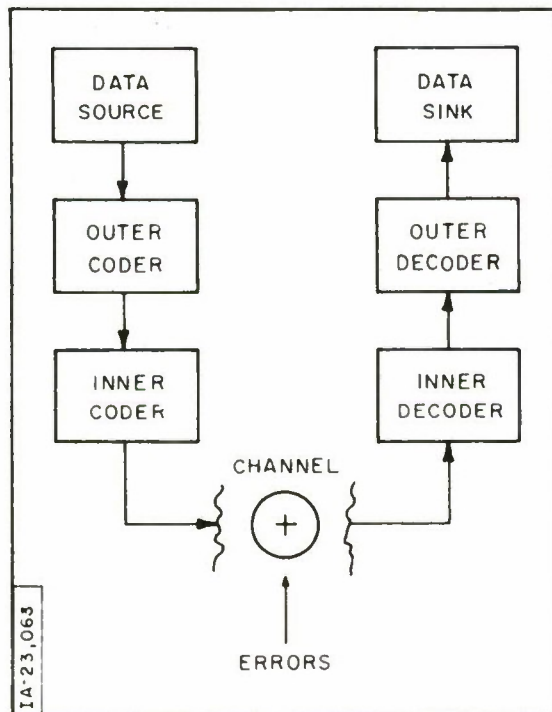
Figure 9. Tandem Code Configuration

SECTION II

JUSTIFICATION FOR THE TANDEM CODING APPROACH

The results presented in previous papers[1,2] and the results of Part I of this paper give a good indication of the capability of cyclic codes, with and without interleaving, in terms of error correction performance. An example of this performance is given in Table II for all the channel data collected with the two modems described in Part I.

Table II

Code Performance on Large Data Samples

	TE-216 Modem	AN/USC-12 Modem
Channel Data Rate	2400 bits/sec	2400 bits/sec
Total Bits	4.8×10^7	1.4×10^8
Total Errors	2.6×10^5	2.5×10^5
Error Rate	5.4×10^{-3}	1.7×10^{-3}
Output Error Rate (24, 12, 3) Code; m = 211	9.3×10^{-5}	4.7×10^{-6}
Output Error Rate (16, 4, 3) Code; m = 211	1.6×10^{-5}	2.1×10^{-6}

Using the formulas of Part I, the overall delay with the $(24, 12, 3)$ code is

$$\frac{2 \, mk}{\text{Information Data Rate}} = \frac{2 \cdot 211 \cdot 12}{1200} = 4.22 \text{ seconds,}$$

and the delay with the $(16, 4, 3)$ code is 2.81 seconds. The overall improvement factors for the two codes using the TE-216 modem and the AN/USC-12 modem are 58 and 337, and 361 and 809, respectively.

It remains, however, to find a better approach. The first indication of such an approach is in the previous papers[1,2] where it is demonstrated that error correction performance increases rapidly as a function of m until $m \approx 89$ and thereafter increases slowly. This is demonstrated graphically in Figure 10 where a set of half-rate codes are evaluated against the test data (previously described in detail[1]). As the interleaving increases beyond M=89 it becomes desirable to find another fast-rising performance curve rather than to continue with one that shows no additional improvement.
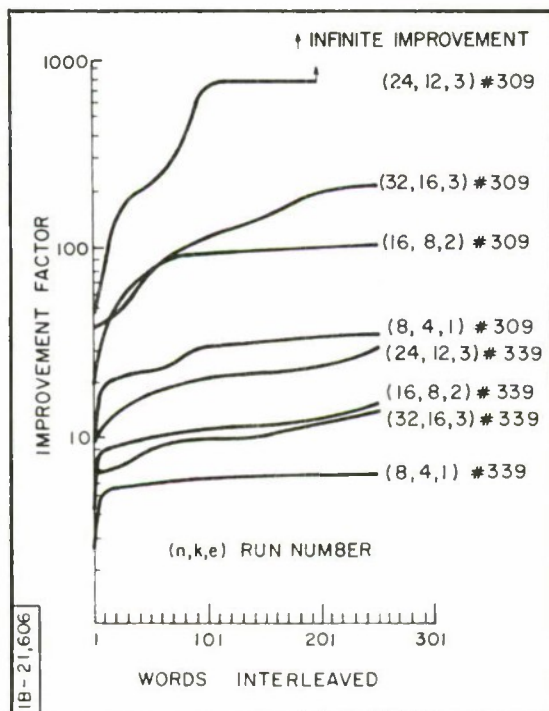
Figure 10. Performance of Half Rate Codes on Typical TE-216 Data

It is demonstrated in Figures 11 and 12, where the consecutive-error and gap functions[1] for test runs 309 and 339* after decoding are presented with m = 1 and m = 89, that when m = 1 the codes correct only the random errors and have no effect on the bursts. The output errors occur in dense bursts with highly pronounced periodic errors (as demonstrated by the vertical jumps in the gap functions). The curves for m = 89, where the code was spread through the data by interleaving, show that the remaining errors are ramdom.[1] Since the errors after decoding are random, it is suggested that, using a tandem coding approach, an inner coder that is interleaved with m ≤ 89 be placed in tandem with an outer coder that will correct the remaining random errors. Before evaluating the tandem coding approach, it is necessary to identify some parameters of the tandem coder.
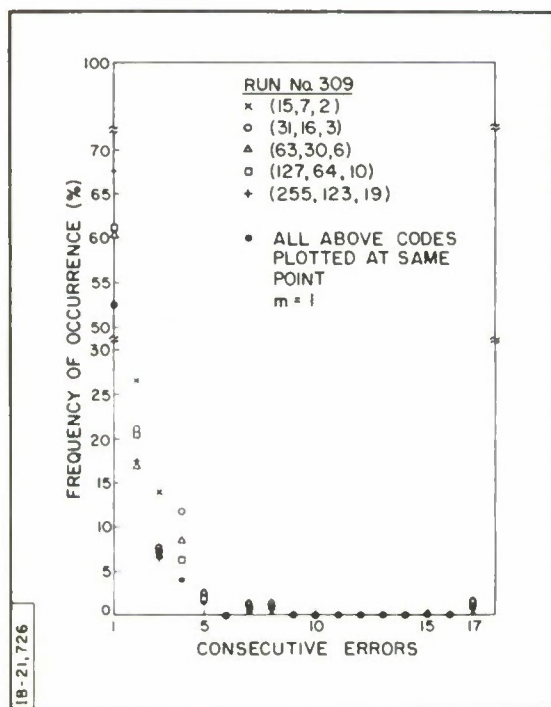
Figure 11a.   Non-Interleaved Decoded Consecutive Error Distribution — Run 309

---

* These test samples are described in detail in the Appendix.
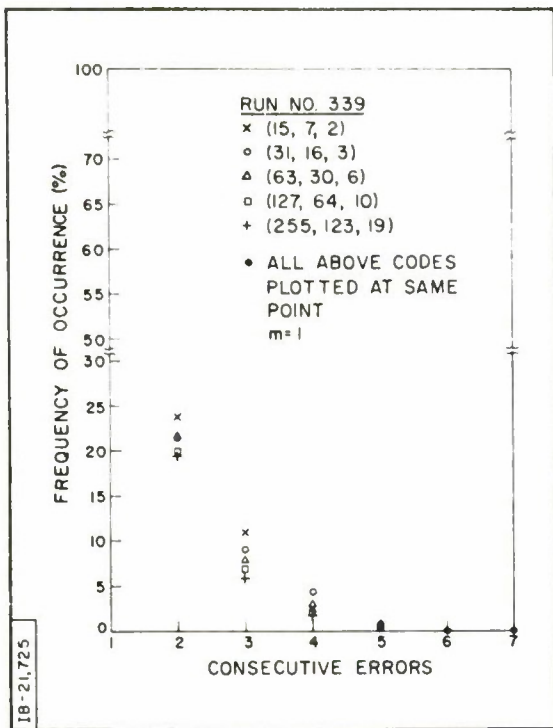
Figure 11b.   Non-Interleaved Decoded
Consecutive Error
Distribution — Run 339

Figure 11c.   Interleaved Decoded
Consecutive Error
Distribution — Runs
309 and 339
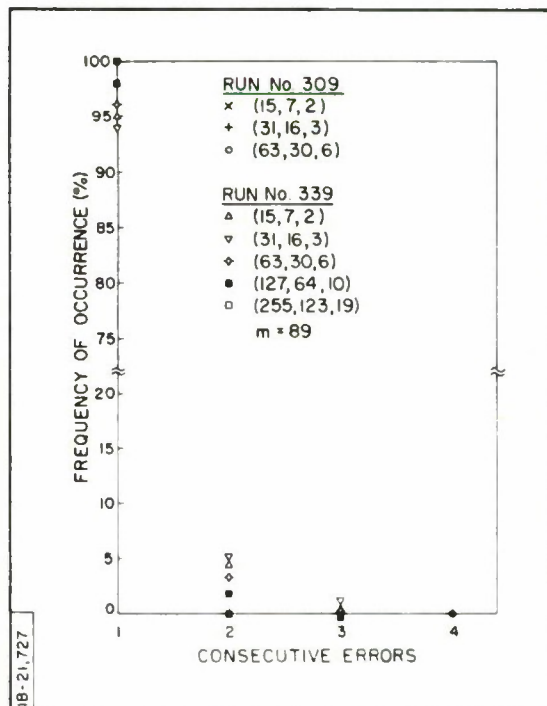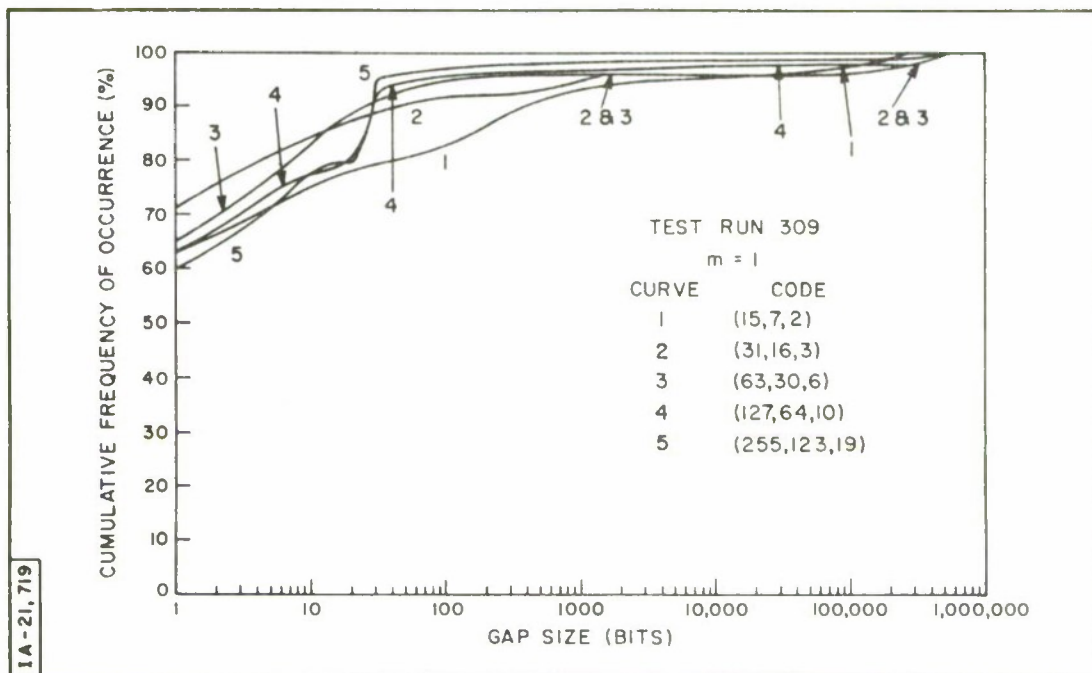
Figure 12a.   Distribution of Gaps After Non-Interleaved Decoding — Run 309
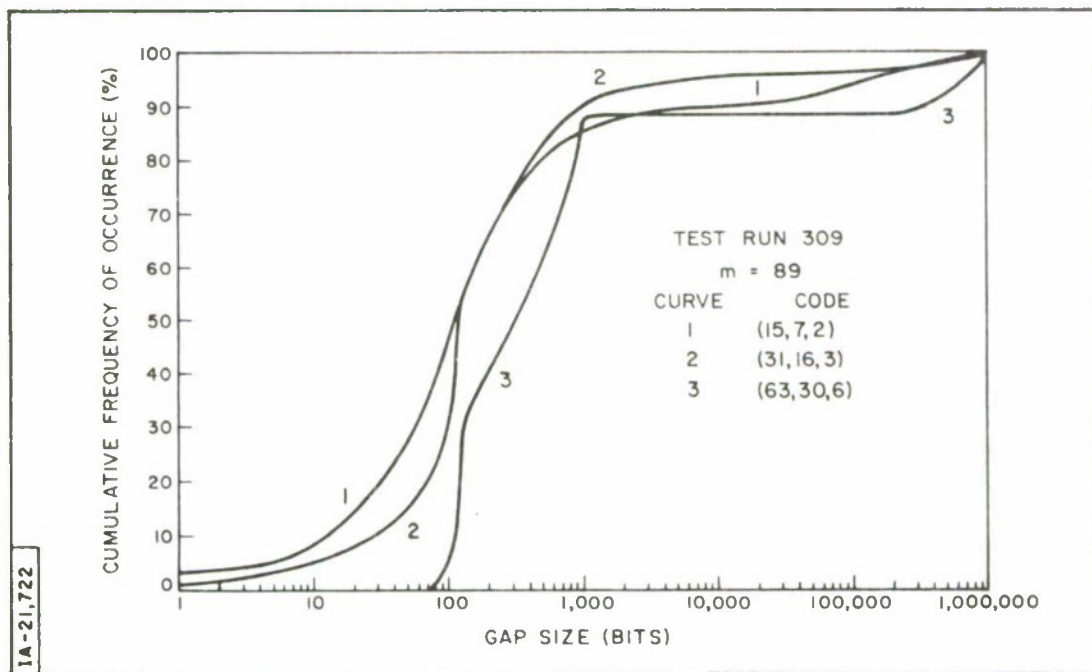
Figure 12b.   Distribution of Gaps After Interleaved Decoding — Run 309

- 37 -
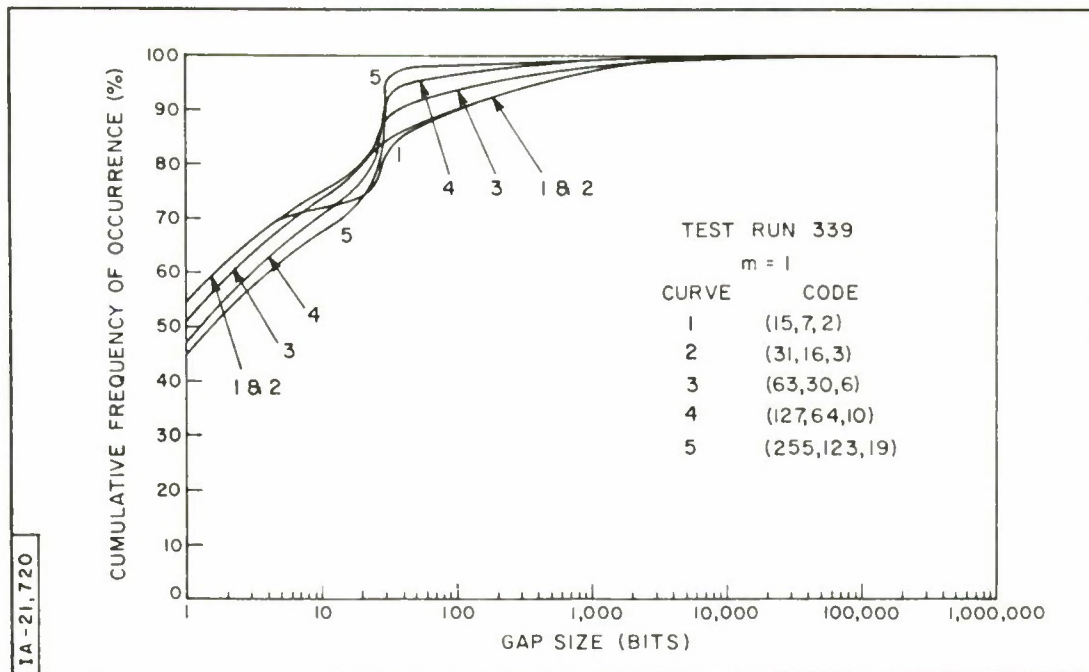
Figure 12c. Distribution of Gaps After Non-Interleaved Decoding – Run 339
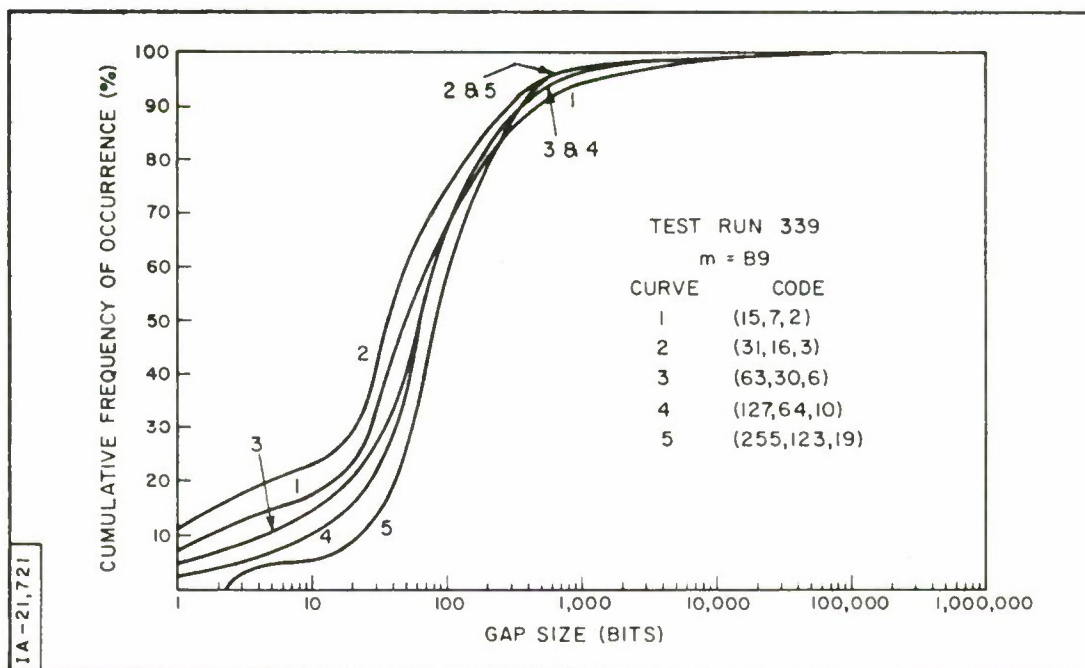
Figure 12d. Distribution of Gaps After Interleaved Decoding – Run 339

- 38 -

SECTION III

TANDEM CODE PARAMETERS

One important parameter of a code is the code rate, which is the ratio of information bits to total bits in a code word. This parameter describes the portion of the communications channel that is being used for information transmission.

The overall code rate (ratio of total information bits to total bits) of this system is the product of the code rates of the two codes. Thus, in order to maintain the same ratio of information to total bits as used for single codes (see Part I), the two codes used here will of necessity have less individual error correction capability than a single code. This is easily verified in terms of Bose-Chaudhuri-Hocquenhem (BCH) codes where,

$$\text{code rate} = \frac{k}{n} = \frac{n-P}{n} \le 1 - \frac{2e+1}{n} \tag{12}$$

If the code rate is increased for a given block length (n), the number of correctable errors (e) must be decreased. This will be balanced out by taking advantage of the additional randomization gained in a decoder output. The delay of the system is the sum of the delays introduced by the two codes. Using the formulas of Part I,

$$\text{Delay} = \frac{2m_o n_o}{(\text{Channel Rate})_o} + \frac{2m_i n_i}{(\text{Channel Rate})_i} \tag{13}$$

where the subscript i refers to the inner decoder and the subscript o refers to the outer decoder. Thus $(\text{Channel Rate})_i$ is the communication channel data rate. Since the number of information bits (k_i) exiting the inner decoder is the number of total bits (n_o) entering the outer decoder,

$$(\text{Channel Rate})_o = \frac{k_i}{n_i} (\text{Channel Rate})_i = \frac{n_o}{n_i} (\text{Channel Rate})_i . \qquad (14)$$

Thus

$$\text{Delay} = \frac{2m_o n_i}{(\text{Channel Rate})_i} + \frac{2m_i n_i}{(\text{Channel Rate})_i} \qquad (15)$$

$$= \frac{2n_i [m_o + m_i]}{(\text{Channel Rate})_i} \qquad (16)$$

where

$\quad m_o$ = number of outer code words interleaved

$\quad m_i$ = number of inner code words interleaved

$\quad n_i$ = total number of bits in a tandem code block

$$n_i = I + P_1 + P_2 \qquad (17)$$

where

$\quad I \quad$ = number of source information bits

$\quad P_1$ = number of parity bits due to outer code

$\quad P_2$ = number of parity bits due to inner code

– 40 –

SECTION IV

PERFORMANCE OF TANDEM CODING

OBJECTIVES OF PERFORMANCE EVALUATION

In this section the performance of tandem coding on actual channel error patterns will be evaluated. The evaluation will be in three parts: (1) the comparison of tandem coding with single coding (only one code) for three of the four typical test samples described in Part I of this paper; (2) a demonstration of the capabilities of tandem codes on a large sample of data; and (3) the performance as a function of delay and code rate. The performance of the modified Golay code, which is the best[1] half-rate code, will be given for reference.

CHANNELS TO BE USED IN EVALUATION

The communication channels to be used in the general evaluation of tandem coding have been previously described[1,3] but will be briefly reviewed here. Exact details on test procedures and communication equipment can be found in the references. Six different modulation systems (modems) were used to collect the data. These are the Kineplex TE-202, Kineplex TE-216, AN/FGC-60, AN/FGC-61A, SC-302, and S-3000X. The TE-216 and AN/FGC-60 were used on a looped basis between Antigua and Ascension, with both transmission and reception at Antigua and rerouting at Ascension. This test was conducted in September and October of 1965. The other four modems were used between Pretoria, South Africa, and Riverhead, Long Island, New York, in the spring of 1964 for a six-week period. The transmission was on a one-way basis with reception at Riverhead. The tests were conducted using the normal operating equipment and circuits on a non-

interference basis. All reception used dual space diversity with rhombic antennas. All test runs were approximately 10 minutes long spaced around the clock.

The test procedure was to continuously transmit a cyclic, repeating, 52-bit digital message from the Transmitter Facility to the Receiver Facility, where the message was detected. This received test message was compared with the original test message suitably delayed to match the total transmission time of the test link. The comparison was made using a modulo-two adder which summed the original test message (delayed) and the received test message. The output of the modulo-two adder indicated a binary 'one' state whenever the two input signals were not the same. The output signal of the modulo-two adder was recorded in real-time on magnetic tape.

The error patterns recorded in the field on magnetic tape were processed through a tape converter which generates an IBM-compatible tape.

An overall description of the data is presented in Table III. Additional information on modem characteristics is available from manufacturers' catalogs.

PERFORMANCE EVALUATION OF TANDEM CODING

Tandem versus Single Coding on Typical Test Samples

In Figure 13 comparative curves of the total delay as a function of code rate for both single and tandem codes are presented for runs 309, 339, and 12 (see Appendix for a description of the test runs) for the case of 100 percent correction. The points for single codes are chosen from Part I by selecting the code that had the lowest delay for a given code rate of all the codes that exhibited 100 percent correction. The tandem codes were evaluated by selecting as an inner code one of the BCH codes listed by Peterson[11]

Table III

Description of Channel Data

Modem	Modulation Technique	Data Rate (bits/sec)	Total Test Time (hrs.)	Total Bits	Average Bit Error Rate
AN/FGC-61A	Frequency-shift keying over 16 tones (only 12 tones used)	900	14.1	4.6×10^7	3.14×10^{-3}
SC-302	Phase-shift keying- frequency differential reference	750	31.1	8.4×10^7	2.17×10^{-3}
TE-202	Four-phase time dif- ferential phase-shift keying	1200	22.6	9.8×10^7	3.48×10^{-3}
S-3000X	Quadraphase, time/ frequency differential coherent phase-shift keying	750	14.0	3.8×10^7	1.78×10^{-3}
AN/FGC-60	16-tone frequency-shift keying	600 1200	11.5 6.0	2.5×10^7 2.6×10^7	1.26×10^{-2} 1.10×10^{-2}
TE-216	Four-phase time dif- ferential phase-shift keying	1200 2400	19.4 6.0	8.4×10^7 5.2×10^7	5.21×10^{-3} 5.16×10^{-3}
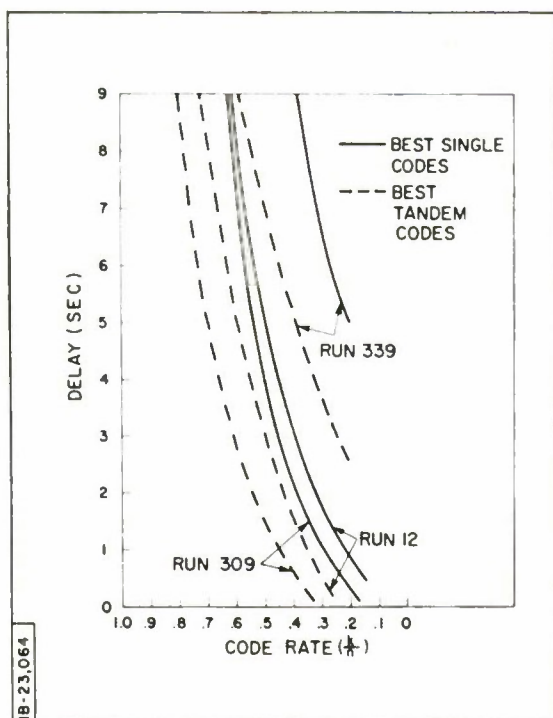
Figure 13. Comparison of Tandem and Single Coding

for block length ℓ, ($15 \leq \ell \leq 255$), interleaving the code by m words (m $\leq$ 89), and decoding. The outer code was similarly selected, and the process was repeated for all ℓ and m and all BCH codes listed by Peterson[11] of rate greater than 0.4 (allowing the evaluation of multiple codes of rate greater than 0.16). From these codes, those that achieved 100 percent correction with the least delay for a given code rate were selected. It is seen in Figure 13 that this approach will, for 100 percent error correction, reduce the system delay by a factor greater than 2 as compared to the delay for single coding (see Part I). More interesting is the fact that test run 339 — which has an error rate of 1.2×10^{-2} and for which it has never before been possible to obtain an improvement factor greater than 100 independent of delay at half-rate[1,2] — is completely corrected, using this approach, with a 7-second delay.

– 44 –

The codes that performed best for each of the three typical data runs were consistent with the results of Part I. The best inner code was always the code which gave the most improvement for a given delay and value of m when considered as a single code, but all inner codes acted as a function of the mn product alone. Since the output errors of the inner decoder were random, the longest cyclic code (n = 255) when used as an outer decoder, performed best (as expected) on the remaining errors. The points in Figure 13 are thus the result of various combinations of all the BCH codes considered used as inner codes and the 255-bit cyclic code, and no interleaving was necessary with the 255-bit outer code to correct the remaining errors.

Performance of Tandem Coding on the Measured HF Channels

To evaluate the performance of a communications channel, a parameter commonly used is the percent of time that the channel average bit error rate (BER) is less than or equal to a given value. For convenience the channel data has been subdivided into equal intervals. Since the data used here had been collected in 10-minute channel intervals, this interval length was selected for use in the simulation. In order that the information bits in a decoded interval will correspond on a one-to-one basis with those in a channel interval, the product of the code rate and the channel interval length is selected as the decoded interval length.

The effect of the use of tandem coding can best be understood by examining the cumulative performance of various channels, using the tandem approach and using the single interleaved code approach, for a given code rate. Half-rate coding, which is generally used in practice, [2] is examined first. The single code used is the modified Golay code of length 24 bits, which corrects 3 errors at rate 0.5 interleaved by m = 125 for a 2.5-second delay in a channel where the information rate is 1200 bits/sec. The tandem code used has as an inner

- 45 -

code a 63-bit length and corrects 3 errors at rate 0.714. The outer 255-bit code corrects 10 errors at rate 0.701. The overall code rate is 0.5005. The inner code is interleaved by 34 code words ($m_i = 34$); the outer code is non-interleaved. The total delay of the tandem code is 2.56 seconds. Using the codes selected, the performance of the best single code (Golay)[1,2] can be compared with the performance of tandem coding while keeping the other parameters constant (code rate, delay).

The results of applying the previously indicated codes to the channel data are presented in Figures 14 through 17. It is evident from these figures that the tandem-coding approach is superior to single encoding. For example, in the channel with the TE-216 modem at 2400 bits/sec (Figures 14 and 15) the channel exhibits error rates poorer than 10^{-4} for 80 percent of the intervals. With the Golay code this occurs for 22 percent of the intervals and with tandem coding it never occurs. Further, 83 percent of the intervals are error-free after tandem decoding, whereas with the Golay code only 52 percent are error-free, and in the actual channel only 2 percent are error-free.

Similar curves for the AN/FGC-60 data are presented in Figures 16 and 17. In Figures 18 and 19 the performance of the Golay code as a function of delay for the AN/FGC-60 1200 bit/sec data is presented, and the same performance curves are given in Figure 20 using the tandem code set previously defined. No curves can be presented for the AN/FGC-60 data at 600 bits/sec, tandem-coded, since the increase of either m_o or m_i yields a channel with no output errors. The same results occur with the TE-216 data. The apparent inconsistency in stating that all of the TE-216 data is corrected with less than 3 seconds delay, while test sample 339 (a member of the sample) required 7 seconds for 100 percent correction, is explained by the fact that 100 percent

Figure 14.   Cumulative Performance Curves, TE-216 Modem — Tandem Coding

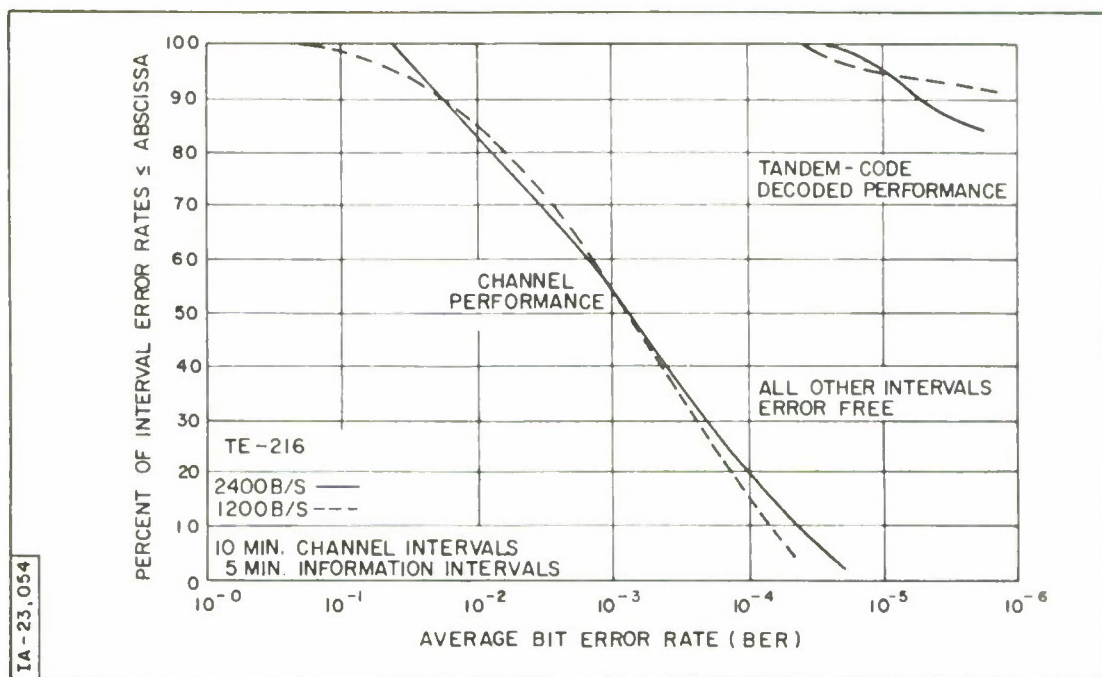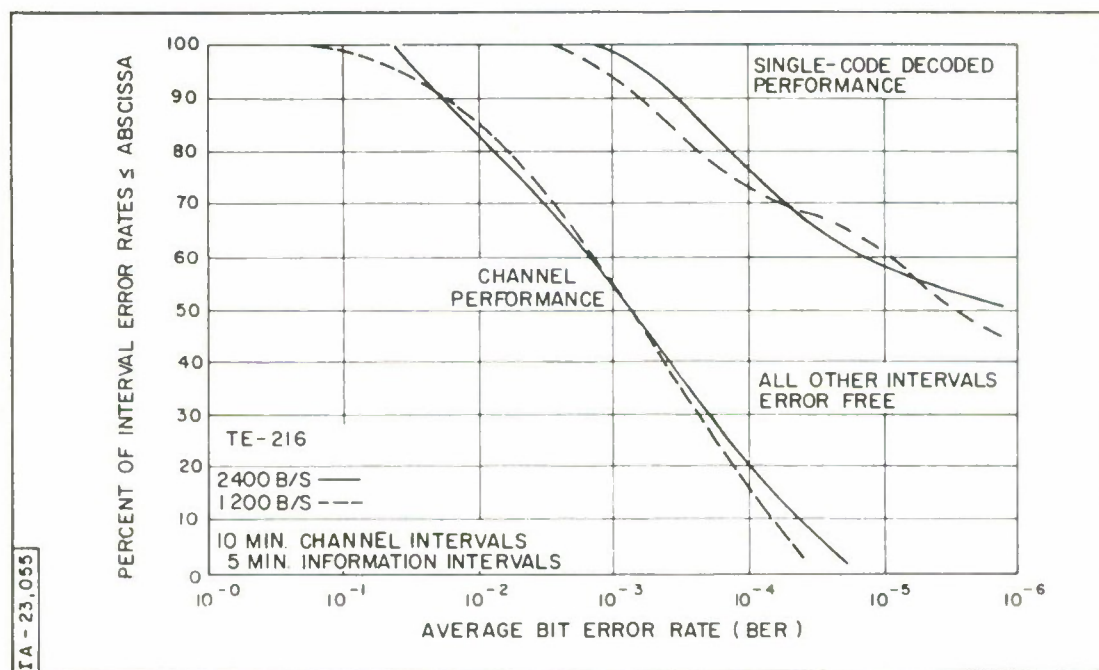Figure 15.   Cumulative Performance Curves, TE-216 Modem — Single Coding

- 47 -
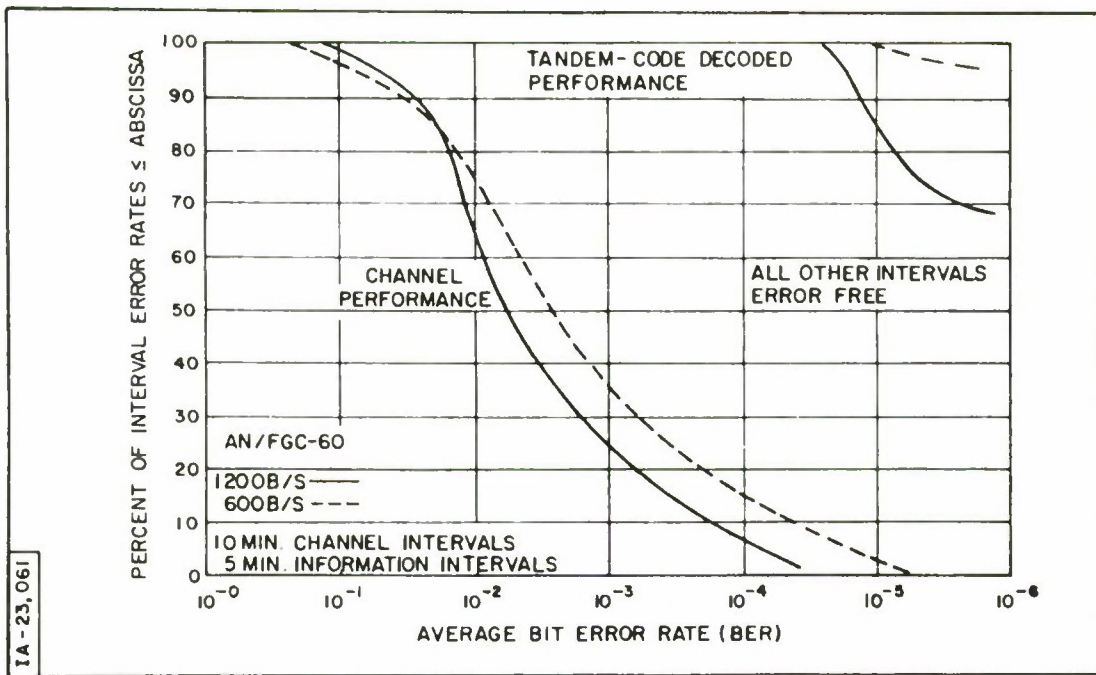
Figure 16.   Cumulative Performance Curves, AN/FGC-60 Modem — Tandem Coding
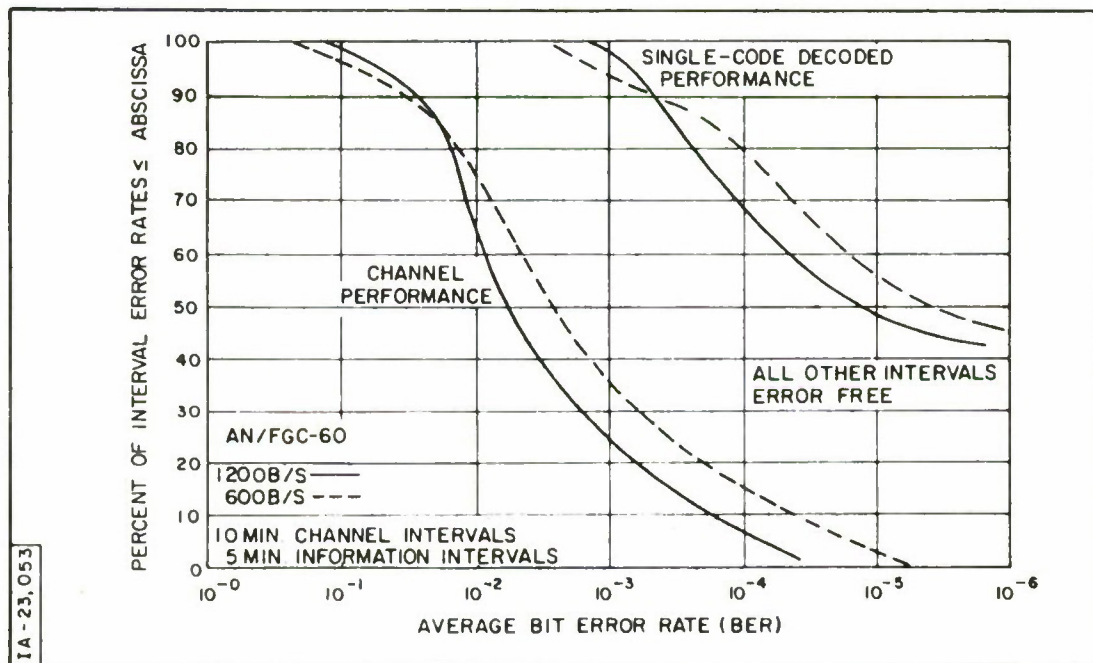
Figure 17.   Cumulative Performance Curves, AN/FGC-60 Modem — Single Coding
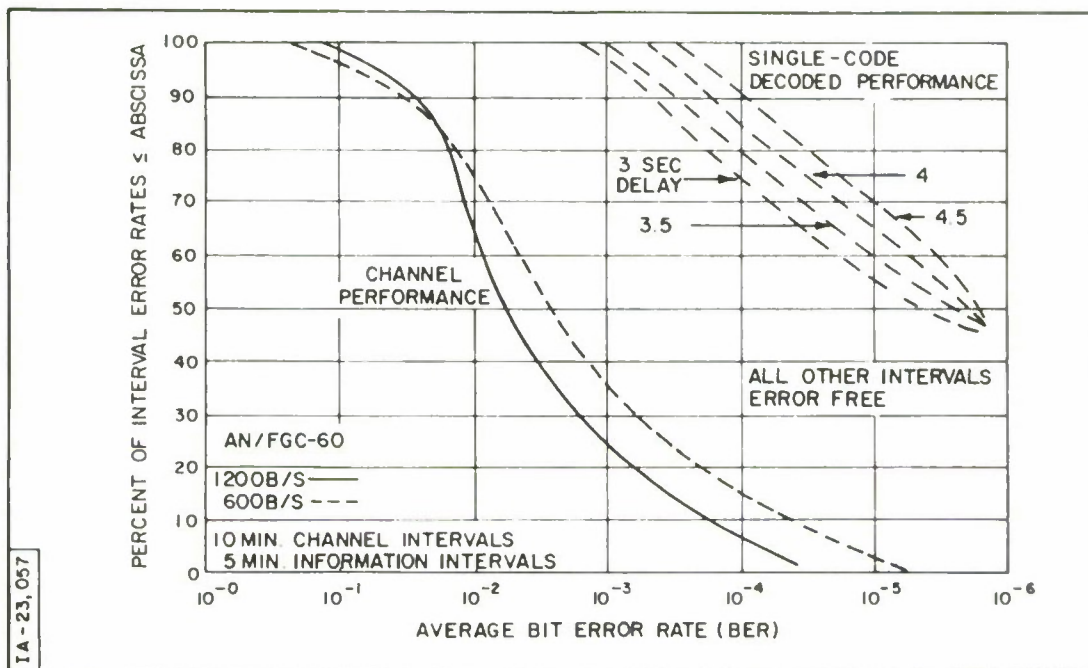
Figure 18.   Single-Code Performance as a Function of Delay in AN/FGC-60,
600 bits/sec Channel
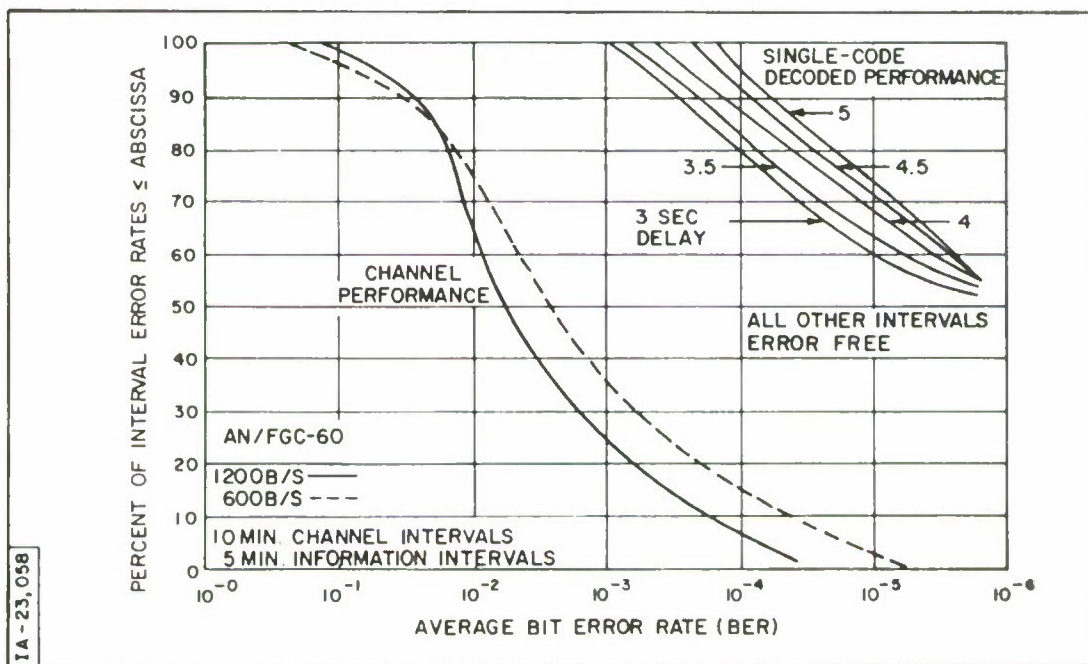
Figure 19.   Single-Code Performance as o Function of Delay in AN/FGC-60,
1200 bits/sec Channel
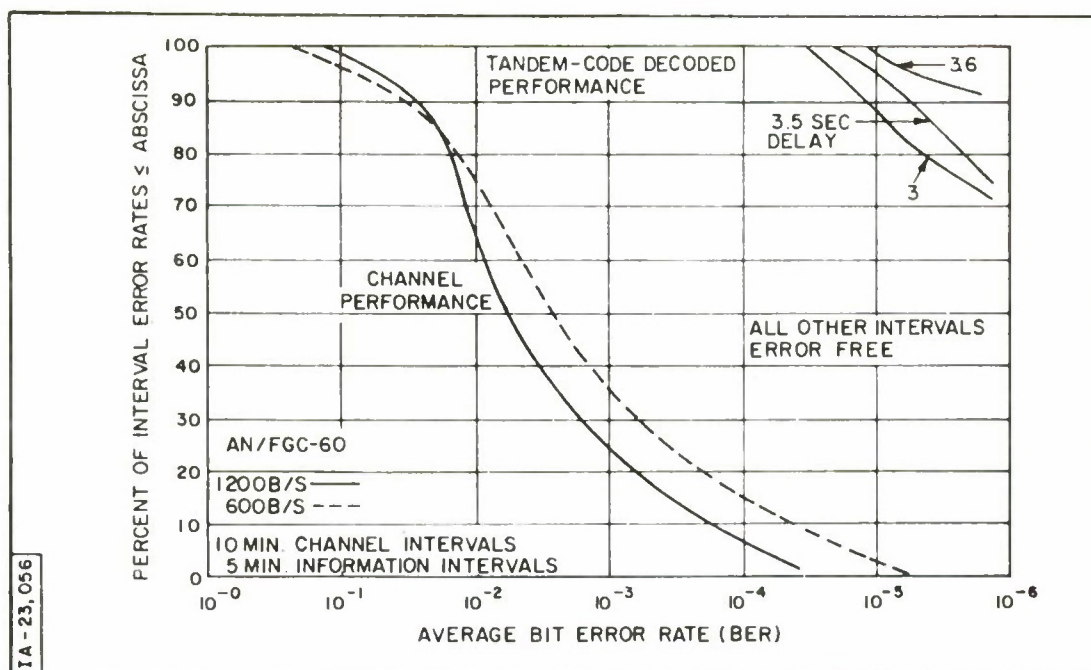
Figure 20.   Tandem Code Performance as a Function of Delay in AN/FGC-60, 1200 bits/sec Channel

correction means correction of all errors in the parity as well as the information, whereas in the large sample evaluation, normal decoding procedures were used which allow for residual errors in the parity, since the parity is discarded after decoding.

The performance of the same tandem code set and of the Golay code, using the other four modems (S-3000X, TE-202, SC-302, and AN/FGC-61A), is presented in Figures 21 and 22.   It is again noted that the tandem code performance is better than the Golay, which from previous evaluation is the best single code.   It should also be noted that with the S-3000X modem, which is similar to the AN/GSC-10,[12] little improvement can be achieved with either technique for the delay of 2.5 seconds.   The reason is that errors occur in long, dense bursts with this modem which can only be corrected with $m \geq 3000$ for the Golay code and $m_i + m_o \geq 300$ for the tandem code considered.   These values represent delays on the order of one minute.

- 50 -
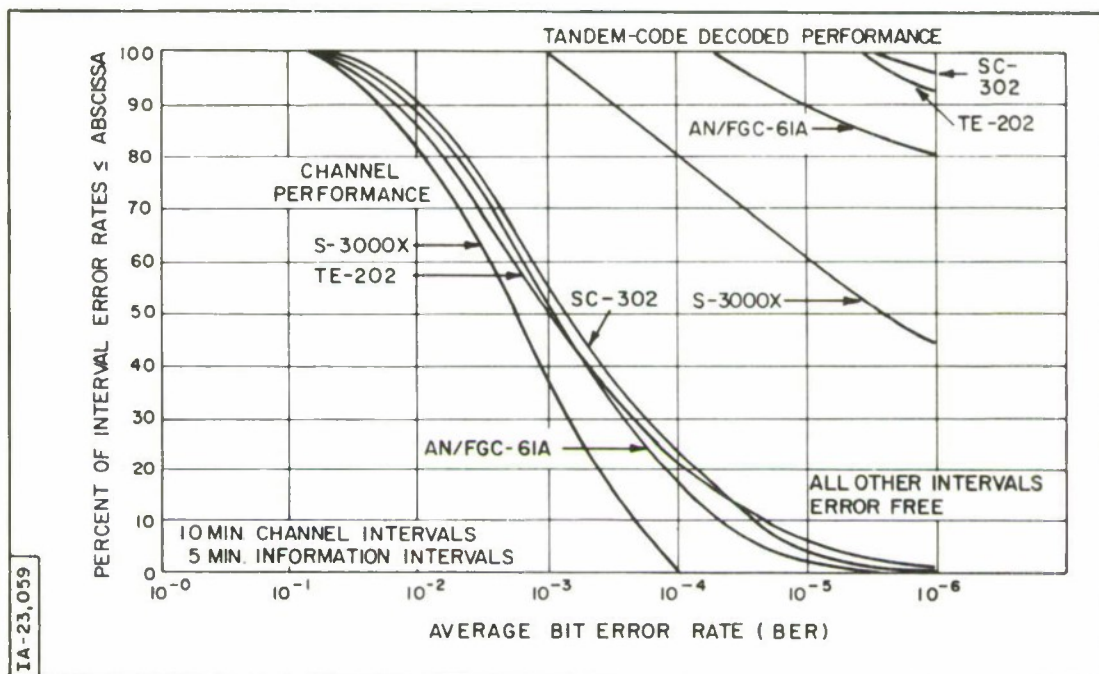
Figure 21. Performance of Tandem Coding on Four Channels
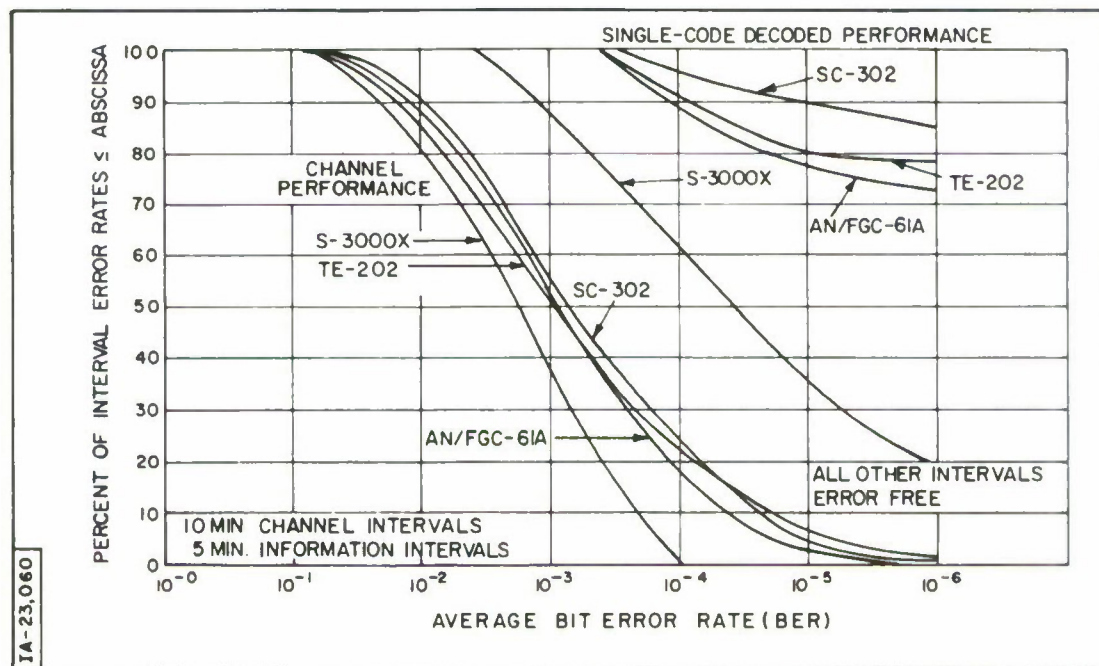
Figure 22. Performance of Single Coding on Four Channels

- 51 -

Performance of Tandem Coding As a Function of Code Rate and Delay

The remaining step is to evaluate tandem coding as a function of code rate. This presents problems, both because there are large numbers of codes to choose from and because delay, being a function of n_i, will be inter-related to code rate, making it difficult to compare codes on the basis of a common delay. Performance curves are not presented, since they are meaningful only in terms of the data used. Instead, when it was discovered that performance was linear in all variables, an empirical relationship between the average interval bit error rate and the percent of interval error rates below the average was determined as a function of delay and code rate.

For the S-3000X, SC-302, TE-202, AN/FGC-61A, AN/FGC-60 and TE-216 at 1200 bits/sec, the empirical relationship was found to be of the form

$$Y = (a + a_R + a_D) X + b + b_R + b_D \tag{18}$$

where

Y = percent of interval error rates $\leq$ abscissa

X = Bit Error Rate = abscissa

a = fixed component of slope factor

b = fixed component of offset

a_R = f (code rate)

a_D = f (delay)

b_R = f (code rate)

b_D = f (delay)

As stated previously a_R, a_D, b_R, and b_D are, in general, functions of both code rate and delay since code rate and delay are in themselves interrelated; however, for the test data collected with the five indicated modems, the effects of code rate and delay were themselves separable into linear relationships. The results of this empirical analysis are tabulated in Table IV. For the TE-216 modem at 2400 bits/sec there was a wide scattering of points for which no relationship in the form of Equation (18) could be found; however, it was found that the relationship for the TE-202 usually gives performance poorer than the

Table IV

Empirical Performance Relations

Data Classification	Relationship
S-3000X	$Y = (7509D - 45054R + 55055)X$ $+ (7.49D - 44.95R + 44.94)$
AN/FGC-61A	$Y = (66360D - 398162R + 408163)X$ $+ (13.25D - 79.60R + 79.59)$
TE-202	$Y = (25544D - 153264R + 163265)X$ $+ (15.30D - 91.84R + 91.83)$
SC-302	$Y = (414999D - 2489999R + 2599000)X$ $+ (15.41D - 92.5R + 92.49)$
AN/FGC-60 1200 bits/sec	$Y = (98710D - 592262R + 594262)X$ $+ (11.71D - 70.28R + 70.25)$
AN/FGC-60 600 bits/sec	$Y = (98710D - 442444R + 444444)X$ $+ (15.85D - 95.11R + 95.05)$
TE-216 1200 bits/sec	$Y = (19274D - 115647R + 117647)X$ $+ (15.29D - 91.76R + 91.74)$

Restrictions

$0 \leq Y \leq 100$

$1 \times 10^{-6} \leq X \leq 1 \times 10^{-2}$

$0.5 < R = \text{code rate} < 1$

$0 < D = \text{delay} \leq 2.5 \text{ sec}$

TE-216 at 2400 bits/sec, and it is recommended that the TE-202 curves be
used as a worst-case estimate. A set of these curves for AN/FGC-61A data
is presented in Figure 23. A subset of the BCH codes listed by Peterson[11]
which were found to give the best actual performance as inner codes is tabu-
lated in Table V. Additionally, comments on the selection of outer codes are

Table V

Best BCH Codes in Tandem Code Evaluation
(by practical evaluation)

Inner Codes			
n	k	e	
7	4	1	
15	11	1	
24	12	3	
31	21	2	
63	51	2	
63	45	3	
63	39	4	
63	36	5	Notes:
127	106	3	
127	99	4	1. $m_i \leq 89$ should be used with inner codes.
127	85	6	
127	71	9	2. Output errors of inner decoder
255	239	2	are sufficiently random that the
255	231	3	outer code can be chosen on delay
255	223	4	and code rate considerations only.
255	207	6	
255	199	7	3. Symbol codes can be used as inner
255	187	9	codes only if the outer code is
255	171	11	interleaved ($m_o > 1$).
255	155	13	
255	131	18	

Figure 23. Empirical Relationship for Code Performance on AN/FGC-61A Data

given as notes on that table. It should be noted that symbol codes[9] which in Part I of this paper were found inferior when used singly are satisfactory as inner codes provided that the outer code is interleaved. As single codes the error correction capability of the symbol codes could easily be exceeded; thus the codes would fail in bursts. In tandem coding the outer interleaved binary code will provide the additional error correction necessary.

SECTION V

CONCLUSION

Interleaving when used in conjunction with a powerful random-error correcting code will provide an improvement factor that is approximately one-tenth of the reciprocal of the input bit error rate. This improvement has been presented graphically in the form of ogive-shaped cumulative distribution curves. For the observed HF channels with bit error rates not worse than 10^{-2}, it is possible through the use of one level of interleaved encoding to correct approximately one-half of the error bursts and leave the remaining errors in a near random state. When the second level of encoding is super-imposed (concatenated) upon the first the result is that the bit error rate at the output of the tandem (concatenated) system is rarely poorer than 10^{-6}. The decoded blocks are almost all error-free provided that the overall code rate is in the neighborhood of one half and the delay is in the range of two to four seconds. Thus it has been demonstrated that the division of redundant bits from a single level of encoding into two concatenated levels of encoding with interleaving is an efficient method of error control in the HF channels considered.

APPENDIX

The probability distribution functions that describe the data presented here are:

1. Distribution of Consecutive Errors
2. Distribution of Gaps
3. Distribution of Burst Lengths
4. Distribution of Burst Densities
5. Distribution of Interval Lengths
6. Distribution of Interval Densities
7. Distribution of Guard Space

In the case of independent random errors the cumulative distributions of consecutive errors and gaps are defined by the following equations:

$$P\left\{e^r c \mid c\right\} = \sum_{j=0}^{r} p^j(1-p) \qquad (19)$$

$$P\left\{c^r e \mid e\right\} = \sum_{j=0}^{r} (1-p)^j p \qquad (20)$$

where

c represents a correct bit

e represents an error bit

r represents the number of consecutive bits

p is the probability of bit error

Relative distributions are constructed using the discrete values of j. The relative distributions of consecutive errors are presented in Figure 24. If the values of p (assuming p equals average bit error rate) for the various runs were inserted into Equation (19), it would be found that in all cases single errors should occur over 98 percent of the time. Only for run 12 does this happen, giving the first indication that run 12 contains random errors. Test run 24 indicates a high occurrence of short dense bursts of errors, and the other runs lie somewhere in between. The same conclusions can be drawn from the gap distribution functions (Figure 25). Test run 12 has a distribution of gaps near that for independent random errors while run 24 demonstrates a high occurrence of short gaps between errors, indicating dense bursts. Test runs 12, 309, and 339 also indicate the occurrence of periodic errors, as demonstrated by the high occurrence of gaps of a specific size.[1] Thus the initial indication is that run 12 exhibits independent random errors and periodic errors, run 24 exhibits dense bursts, and runs 309 and 339 have periodic errors and lie between 12 and 24 in burstiness.

The above discussion of probability functions does not present the complete picture. There is no information about the length of bursts, nor is there information relative to the interval between bursts (guard space).

A burst is defined as a region of the serial data stream where the following properties hold: A minimum number of errors, M_e, are contained in the region and the minimum density of errors in the region is Δ. Both of these conditions for the chosen values of M_e and Δ must be satisfied for the region to be defined as a burst. The density of errors is defined as the ratio of bits in error to the total number of bits in the region.

The burst probability density function is defined as the probability of occurrence of a burst of size N where N is any positive integer. The burst
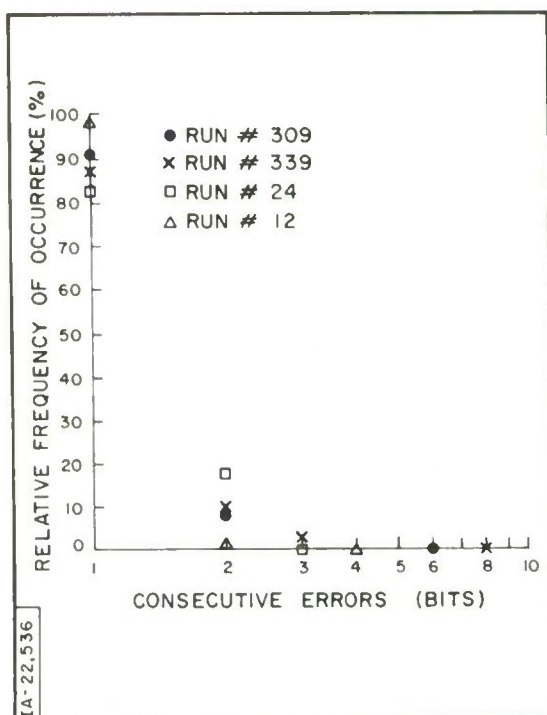
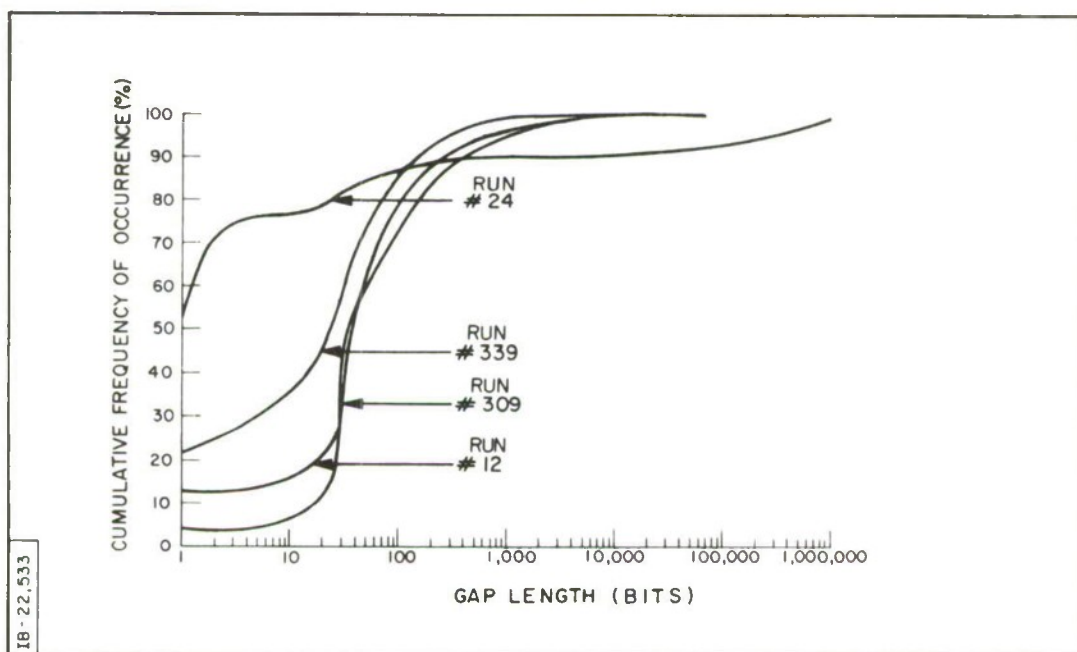Figure 24. Relative Frequency of
Consecutive Errors

Figure 25. Distribution of Gaps Between Errors

size is measured in terms of the total number of bits in the burst. A separate burst probability density function may be determined for each pair of values of Δ and M_e.

The following properties hold for the burst: The burst always begins with a bit in error and ends with a bit in error; a burst may contain correct bits; each burst is immediately preceded and followed by an interval in which the density of errors is less than Δ. The minimum number of errors (M_e) in a burst has been chosen to be two (2) for all the data included here.

The interval is defined as the region of the serial data stream where the following properties hold: The minimum density of errors is less than Δ, and the region begins and ends in a correct bit; an interval may contain errors; an interval is always immediately preceded and followed by a burst. Thus, each and every bit in the data stream must lie in either a burst region or an interval region.

The interval probability density function is defined as the probability of occurrence of an interval of length L, where L is any positive integer. The interval probability density is a joint function of both Δ and M_e.

The guard space ratio is defined as the ratio of the interval length to burst length preceding it.

The burst length and density distributions are presented in Figures 26 and 27. Examination of these figures verifies that run 24 is composed of short dense bursts. The distributions for test run 12, however, are similar to those for runs 309 and 339. Thus, it is concluded that the random and periodic errors of run 12 are actually contained within bursts.

From Figures 28, 29, and 30 it is seen that the intervals between the bursts are generally long, contain few errors, and in the case of run 24
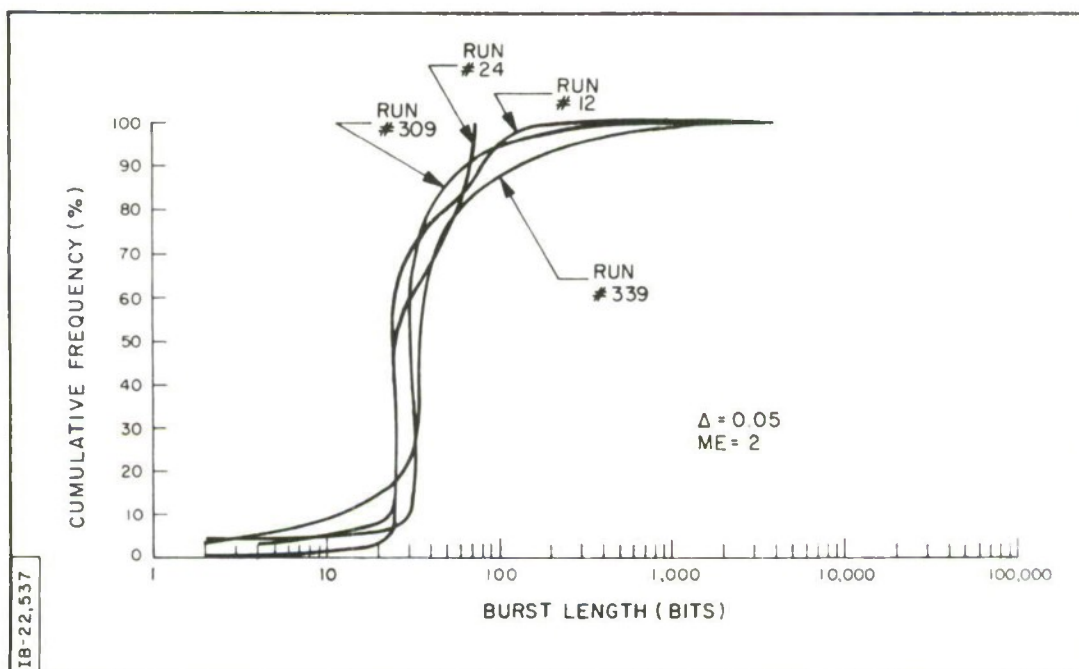
Figure 26.   Distribution of Lengths of Bursts
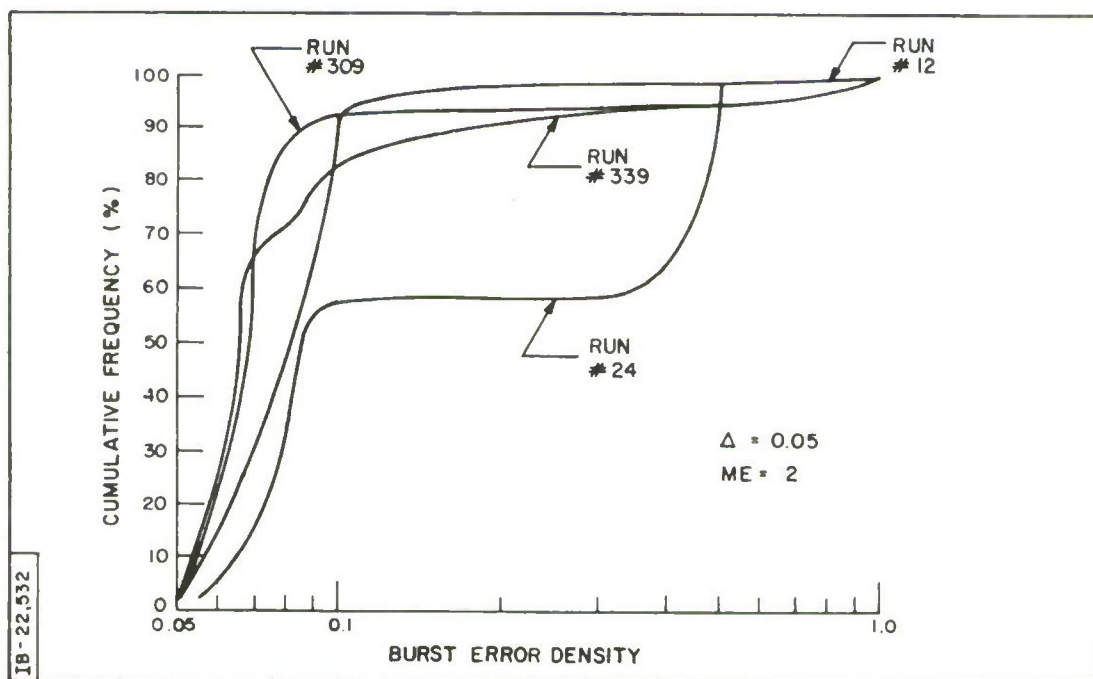
Figure 27.   Distribution of Burst Error Densities
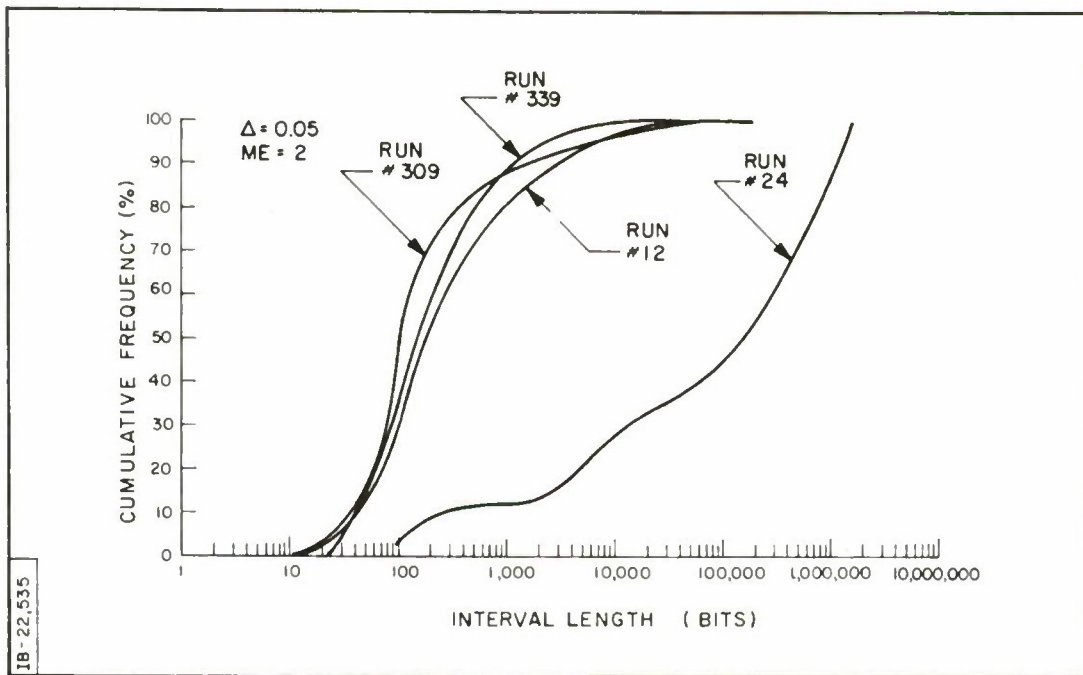
Figure 28. Distribution of Lengths of Intervals

Figure 29. Distribution of Interval Error Densities

- 64 -

Figure 30. Distribution of Guard Space

provide excellent guard space protection. The guard space protection in runs 12, 339, and 309 is insufficient in that respectively over 10 percent, 18 percent, and 26 percent of the bursts are followed by intervals that are actually shorter than the burst (guard space ratio < 1).

REFERENCES

1. K. Brayer, O. Cardinale, "Evaluation of Error Correction Block Encoding for High Speed HF Data," IEEE Trans. on Communication Technology, June 1967.
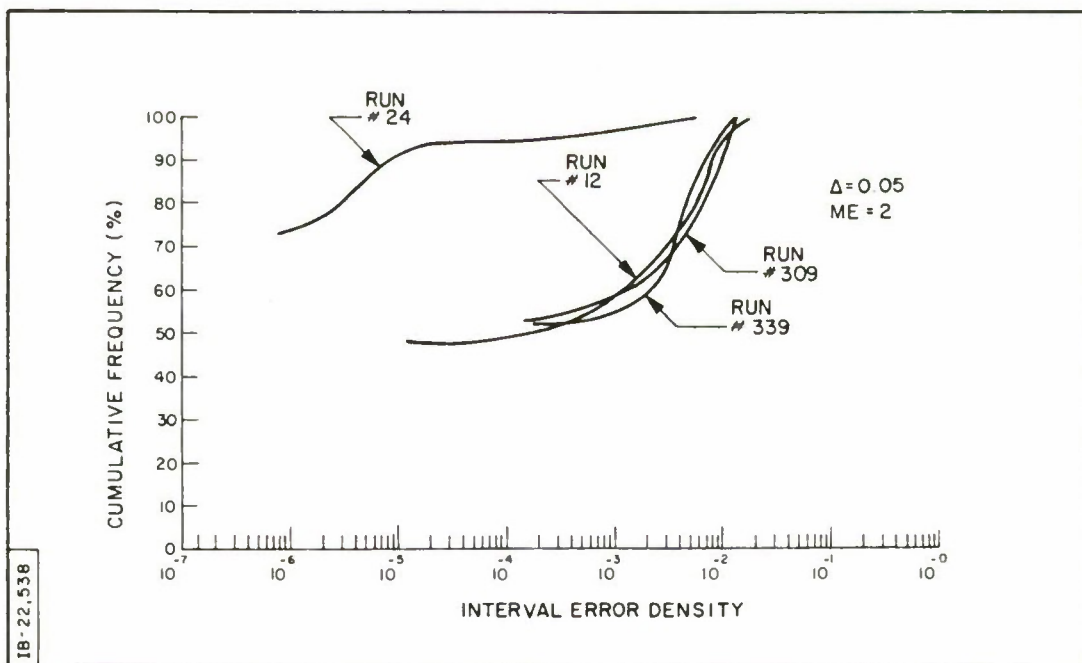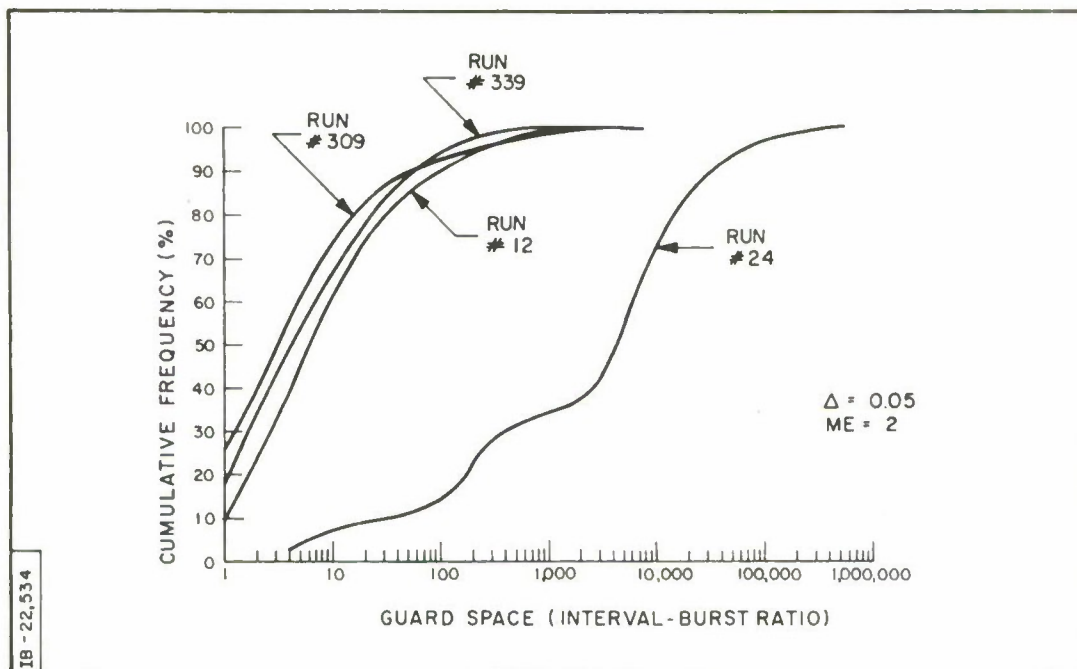
2. K. Brayer, Design Considerations for Forward Error Control Equipment for HF Radio, ESD-TR-67-356, August 1967.

3. K. Brayer, Error Patterns Measured on Transequatorial HF Communication Links, MTP-46, The MITRE Corporation, Bedford, Mass., December 1966. (To be published, IEEE Trans. on Communication Technology, April 1968)

4. R. C. Bose, D. K. Ray-Chaudhuri, "On a Class of Error Correcting Binary Group Codes," Inf. and Control, 3, 68-79 (1960).

5. R. C. Bose, D. K. Ray-Chaudhuri, "Further Results on Error Correcting Binary Group Codes," Inf. and Control, 3, 279-290 (1960).

6. K. Brayer, O. Cardinale, HF Channel Error Statistics Description (I), ESD-TR-66-290, June 1966.

7. K. Brayer, O. Cardinale, HF Channel Error Statistics Description (II), ESD-TR-66-107, May 1966.

8. D. Gorenstein, N. Zierler, "A Class of Error Correcting Codes in P^M Symbols," J. Siam, 9, June 1961.

9. K. Brayer, Error Control Techniques Using Binary Symbol Burst Codes, MTP-55, The MITRE Corporation, Bedford, Mass., May 1967. (To be published, IEEE Trans. on Communication Technology, April 1968)

10. G. D. Forney, Jr., Concatenated Codes, TR 440, MIT Research Lab of Electronics, December 1965.

11. W. W. Peterson, Error Correcting Codes, The MIT Press and John Wiley & Sons, Inc., 1961, p. 166.

12. M. S. Zimmerman, A. Kirsch, "The AN/GSC-10 (KATHRYN) Variable Rate Data Modem for HF Radio," IEEE Trans. on Communication Technology, April 1967.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2a. REPORT SECURITY CLASSIFICATION
The MITRE Corporation Bedford, Massachusetts	UNCLASSIFIED
	2b. GROUP N/A

3. REPORT TITLE

THE IMPROVEMENT OF DIGITAL HF COMMUNICATION THROUGH CODING

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
   N/A

5. AUTHOR(S) *(First name, middle initial, last name)*

Kenneth Brayer

6. REPORT DATE May 1968	7a. TOTAL NO. OF PAGES 68	7b. NO. OF REFS 12
8a. CONTRACT OR GRANT NO. AF 19(628)-5165	9a. ORIGINATOR'S REPORT NUMBER(S) ESD-TR-68-128	
b. PROJECT NO. 705B		
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* MTP-75	
d.		

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES N/A	12. SPONSORING MILITARY ACTIVITY Aerospace Instrumentation Program Office, Electronic Systems Division, Air Force Systems Command, L.G. Hanscom Field, Bedford, Massachusetts

13. ABSTRACT

In previous papers the technique of error correction of digital data through the use of interleaved cyclic codes and a set of probability functions for the evaluation of error patterns have been presented. In this paper the previous results are extended to a wide range of BCH and symbol codes. A set of simple equations is presented for the description of an interleaved cyclic code and its associated delay, and a method is presented which allows for a significant increase in error rate improvement at a reduction in the delay time introduced into the channel. It is demonstrated that the performance of interleaved cyclic codes is sufficient to correct all types of measured HF error patterns; that, using delay as a basis of comparison, only the total bit interleaving is important in achieving error correction; and that it is possible to get almost 100 percent error correction for delays under 3 seconds for all channel conditions measured.

DD FORM 1 NOV 65 **1473**

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
SYSTEMS AND MECHANISMS						
Data Transmission Systems						
Multi-Channel Radio System						
Voice Communication (HF) Systems						
INFORMATION THEORY						
Cyclic Coding						
Concatenated Coding						